

TRACE IDEALS FOR PSEUDO-DIFFERENTIAL OPERATORS AND THEIR COMMUTATORS WITH SYMBOLS IN α -MODULATION SPACES

MASAHARU KOBAYASHI, MITSURU SUGIMOTO, AND NAOHITO TOMITA

ABSTRACT. The fact that symbols in the modulation space $M^{1,1}$ generate pseudo-differential operators of the trace class was first mentioned by Feichtinger and the proof was given by Gröchenig [12]. In this paper, we show that the same is true if we replace $M^{1,1}$ by more general α -modulation spaces which include modulation spaces ($\alpha = 0$) and Besov spaces ($\alpha = 1$) as special cases. The result with $\alpha = 0$ corresponds to that of Gröchenig, and the one with $\alpha = 1$ is a new result which states the trace property of the operators with symbols in the Besov space. As an application, we also discuss the trace property of the commutator $[\sigma(X, D), a]$, where $a(x)$ is a Lipschitz function and σ belongs to an α -modulation space.

1. INTRODUCTION

In our previous paper [17], we have discussed the L^2 -boundedness of pseudo-differential operators with symbols in the α -modulation spaces $M_{s,\alpha}^{p,q}$ ($0 \leq \alpha \leq 1$), a parameterized family of function spaces, which include the modulation spaces $M_s^{p,q}$ ($\alpha = 0$) and the Besov spaces $B_s^{p,q}$ ($\alpha = 1$) as special cases. More precisely, the symbol $\sigma \in M_{(an/2, an/2), (\alpha, \alpha)}^{(\infty, \infty), (1, 1)}$, which means $\sigma(x, \xi)$ belongs to $M_{an/2, \alpha}^{\infty, 1}$ in both x and ξ , generates the $L^2(\mathbb{R}^n)$ -bounded pseudo-differential operator. Especially in the case $\alpha = 0$ (resp. $\alpha = 1$), this result corresponds to that of Sjöstrand [23] (resp. Sugimoto [25]), which says the L^2 -boundedness of the operators with symbols in the modulation space $M^{\infty, 1}$ (resp. Besov space $B_{(n/2, n/2)}^{(\infty, \infty), (1, 1)}$).

On the other hand, it is known that symbols in the modulation space $M^{1,1}$ generate pseudo-differential operators of the trace class. This fact was first mentioned by Feichtinger and the proof was given by Gröchenig [12]. As a corollary, we get the result by Daubechies [6] which says that $\sigma \in L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ has the same property

$$\|\sigma(X, D)\|_{\mathcal{I}_1} \leq C (\|\langle x; \xi \rangle^s \sigma(x, \xi)\|_{L^2(\mathbb{R}^{2n})} + \|\langle x; \xi \rangle^s \widehat{\sigma}(x, \xi)\|_{L^2(\mathbb{R}^{2n})})$$

for $s > 2n$, where $\|\cdot\|_{\mathcal{I}_1}$ is the trace norm, $\mathbb{R}^{2n} = \mathbb{R}_x^n \times \mathbb{R}_\xi^n$ and $\langle x; \xi \rangle = (1 + |x|^2 + |\xi|^2)^{1/2}$ (see Gröchenig [13, Corollary 8.38]). Further developments in this direction can be also seen in Cordero-Gröchenig [5], Fernández-Galbis [8], Gröchenig-Heil [14], Labate [18] and Toft [28, 29].

On account of our L^2 -boundedness result, it is natural to expect that the same trace property is true if we replace $M^{1,1}$ by more general α -modulation spaces

2000 *Mathematics Subject Classification.* 42B35, 47B10, 47G30.

Key words and phrases. α -modulation spaces, Besov spaces, trace class, pseudo-differential operators, commutators.

$M_{(\alpha n/2, \alpha n/2), (\alpha, \alpha)}^{(1,1), (1,1)}$. We remark that the notion of α -modulation spaces was introduced by Gröbner [11], and developed by the works of Feichtinger-Gröbner [7], Borup-Nielsen [1, 2] and Fornasier [9]. The precise definition of them will be given later in Section 2. The following is our main theorem:

Theorem 1.1. *Let $0 \leq \alpha \leq 1$. Then there exists a constant $C > 0$ such that*

$$\|\sigma(X, D)\|_{\mathcal{I}_1} \leq C \|\sigma\|_{M_{(\alpha n/2, \alpha n/2), (\alpha, \alpha)}^{(1,1), (1,1)}}$$

for all $\sigma \in M_{(\alpha n/2, \alpha n/2), (\alpha, \alpha)}^{(1,1), (1,1)}(\mathbb{R}^n \times \mathbb{R}^n)$.

Theorem 1.1 with $\alpha = 0$, which requires $\sigma \in M^{1,1}$, is the result by Gröchenig [12, 13]. On the other hand, Theorem 1.1 with $\alpha = 1$ states the trace property of the operators with symbols in the Besov space $B_{(n/2, n/2)}^{(1,1), (1,1)}$, but there seem to be few literature mentioning this fact. We remark that the spaces $M^{1,1}$ and $B_{(n/2, n/2)}^{(1,1), (1,1)}$ have no inclusion relation with each other (see Proposition A.1 in Appendix A). The proof of Theorem 1.1 will be give in Section 3. It follows the same spirit as used in [13], but requires extra arguments. In fact, roughly speaking, modulation spaces are characterized by the uniform decomposition $\{k + [-1, 1]^n\}_{k \in \mathbb{Z}^n}$ while Besov spaces the dyadic one $\{\{\xi \in \mathbb{R}^n : 2^{j-1} \leq |\xi| \leq 2^{j+1}\}\}_{j \geq 1}$. The main obstacle of the proof comes from the non-uniformity of the decomposition used to define the α -modulation spaces, because they are defined by an intermediate type of uniform and dyadic ones. In order to overcome the difficulty, we introduce a modified version of Rihaczek distribution (see Section 3), whose original one was used in [13] and works only for the uniform decomposition.

We mention here the relation between known results and ours. We have already mentioned the result by Daubechies [6] which says that $\sigma \in L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s > 2n$) is sufficient for the corresponding operator to be of the trace class. This result is a direct consequence of the inclusion $L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}) \subset M^{1,1}(\mathbb{R}^{2n})$ ($s > 2n$) (see Proposition A.2 (1)). But there is a significant improvement by Heil-Ramanathan-Topiwala [15] and Gröchenig-Heil [14], which says that $\sigma \in L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s > n$) is sufficient. This result includes the pioneering one

$$\|\sigma(X, D)\|_{\mathcal{I}_1} \leq C \sum_{|\alpha| + \dots + |\beta'| \leq 2k} \|x^\alpha \xi^\beta \partial_x^{\alpha'} \partial_\xi^{\beta'} \sigma(x, \xi)\|_{L^2(\mathbb{R}^{2n})}$$

($2k > n$) by Hörmander [16] (see also Gröchenig [13, Corollary 8.40]). On the other hand, we can say that two conditions $\sigma \in M^{1,1}$ and $\sigma \in L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s > n$) are independent ones since we have $M^{1,1}(\mathbb{R}^{2n}) \not\subset L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s > n$) and $M^{1,1}(\mathbb{R}^{2n}) \not\supset L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s \leq 2n$) (see Proposition A.2 (2), (3)). Furthermore our new condition $\sigma \in B_{(n/2, n/2)}^{(1,1), (1,1)}$ is also independent of them since $B_{(n/2, n/2)}^{(1,1), (1,1)}(\mathbb{R}^n \times \mathbb{R}^n) \not\subset L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ ($s > n$) (see Proposition A.3 (2)). Although we cannot expect the inclusion $B_{(n/2, n/2)}^{(1,1), (1,1)}(\mathbb{R}^n \times \mathbb{R}^n) \supset L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ for $s > n$, it is true at least for $s > 2n$ (see Proposition A.3 (1)), hence Theorem 1.1 with $\alpha = 1$ includes Daubechies' one again.

As an application of Theorem 1.1, we also discuss the trace property of the commutator $[\sigma(X, D), a]$, where $a(x)$ is a Lipschitz function. The L^2 -boundedness of the commutator was discussed by Calderón [3], Coifman-Meyer [4] and Marschall [19], where σ belongs to Hörmander's class $S_{\rho, \delta}^\rho$ ($\delta \leq \rho$, $0 \leq \delta < 1$). In [17], we

have generalized the result with $\rho = \delta = 0$ to the case when $\sigma \in M_{(\alpha n/2, \alpha n+1), (\alpha, \alpha)}^{(\infty, \infty), (1, 1)}$. We can again expect the trace property of the commutator if we assume $\sigma \in M_{(\alpha n/2, \alpha n+1), (\alpha, \alpha)}^{(1, 1), (1, 1)}$ instead, replacing ∞ by 1. In fact we have the following theorem:

Theorem 1.2. *Let $0 \leq \alpha \leq 1$. Then there exists a constant $C > 0$ such that*

$$\|[\sigma(X, D), a]\|_{\mathcal{I}_1} \leq C \|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), (\alpha, \alpha)}^{(1, 1), (1, 1)}}$$

for all Lipschitz functions a and $\sigma \in M_{(\alpha n/2, \alpha n+1), (\alpha, \alpha)}^{(1, 1), (1, 1)}(\mathbb{R}^n \times \mathbb{R}^n)$.

The proof of Theorem 1.2 will be give in Section 4. We finally remark that the result on the Schatten class \mathcal{I}_p can be obtained by interpolation argument. In fact, it is known that $\sigma(X, D)$ is a Hilbert-Schmidt operator if and only if $\sigma \in L^2(\mathbb{R}^{2n})$, and we have $\|\sigma(X, D)\|_{\mathcal{I}_2} = \|\sigma\|_{L^2(\mathbb{R}^{2n})}$ (see Pool [21]). Moreover we can easily see that $\|\sigma\|_{L^2(\mathbb{R}^{2n})} \asymp \|\sigma\|_{M_{(0, 0), (\alpha, \alpha)}^{(2, 2), (2, 2)}}$, that is, $\sigma(x, \xi) \in L^2(\mathbb{R}^{2n})$ if and only if $\sigma(x, \xi)$ belongs to $M_{0, \alpha}^{2, 2}$ in both x and ξ . Hence $\|\sigma(X, D)\|_{\mathcal{I}_2} \asymp \|\sigma\|_{M_{(0, 0), (\alpha, \alpha)}^{(2, 2), (2, 2)}}$, and if we interpolate it with Theorem 1.1, then we have

$$\|\sigma(X, D)\|_{\mathcal{I}_p} \leq C \|\sigma\|_{M_{(\alpha n(1/p-1/2), \alpha n(1/p-1/2)), (\alpha, \alpha)}^{(p, p), (p, p)}}$$

for $1 \leq p \leq 2$. On account of the argument above, we only discuss the trace class \mathcal{I}_1 in this paper.

2. PRELIMINARIES

We first review some of the standard facts on singular values of compact operators, following Zhu [31, Chapter 1] and Simon [22]. Let $1 \leq p < \infty$. The singular values $s_j(T)$ of a compact operator T on $L^2(\mathbb{R}^n)$ are the eigenvalues $\lambda_j(|T|)$ of the positive compact operator $|T| = (T^*T)^{1/2}$, where T^* is the adjoint of T . We say that a compact operator T belongs to the Schatten class \mathcal{I}_p if $\{s_j(T)\}_{j=1}^\infty \in \ell^p$. In this case, we write $T \in \mathcal{I}_p$, and define the norm on \mathcal{I}_p by $\|T\|_{\mathcal{I}_p} = \left(\sum_{j=1}^\infty s_j(T)^p\right)^{1/p}$. In particular, \mathcal{I}_1 and \mathcal{I}_2 are called the trace and Hilbert-Schmidt classes, respectively. It is known that for every $j \in \mathbb{Z}_+ = \{0, 1, 2, \dots\}$

$$s_{j+1}(T) = \inf\{\|T - F\|_{\mathcal{L}(L^2)} : F \in \mathcal{F}_j\},$$

where $\mathcal{L}(L^2(\mathbb{R}^n))$ is the space of all bounded linear operators on $L^2(\mathbb{R}^n)$, and \mathcal{F}_j is the class of all linear operators with rank less than or equal to j ([31, Theorem 1.34 (a)]). Consequently,

$$(2.1) \quad \|T\|_{\mathcal{L}(L^2)} = s_1(T) \leq \|T\|_{\mathcal{I}_p}.$$

Since $\|T\|_{\mathcal{I}_p} = \|T^*\|_{\mathcal{I}_p}$ ([31, p.18]) and

$$s_{j+1}(T) = \min_{f_1, \dots, f_j} \max\{\|Tf\| : \|f\|_{L^2} = 1, \langle f, f_i \rangle = 0, 1 \leq i \leq j\}$$

([31, Theorem 1.34 (b)]), where $\langle \cdot, \cdot \rangle$ denotes the L^2 -inner product, we see that

$$(2.2) \quad \|ST\|_{\mathcal{I}_p} \leq \|S\|_{\mathcal{L}(L^2)} \|T\|_{\mathcal{I}_p} \quad \text{and} \quad \|ST\|_{\mathcal{I}_p} \leq \|S\|_{\mathcal{I}_p} \|T\|_{\mathcal{L}(L^2)}.$$

If $T \in \mathcal{I}_p$, then

$$(2.3) \quad \|T\|_{\mathcal{I}_p} = \sup \left(\sum_{j=1}^\infty |\langle Tf_j, g_j \rangle|^p \right)^{1/p},$$

where the supremum is taken over all orthonormal systems $\{f_j\}, \{g_j\}$ in $L^2(\mathbb{R}^n)$. Conversely, if $T \in \mathcal{L}(L^2(\mathbb{R}^n))$ and the right hand side of (2.3) is finite, then T is a compact operator and $T \in \mathcal{I}_p$ ([22, Proposition 2.6]).

Let $\mathcal{S}(\mathbb{R}^n)$ and $\mathcal{S}'(\mathbb{R}^n)$ be the Schwartz spaces of all rapidly decreasing smooth functions and tempered distributions, respectively. We define the Fourier transform $\mathcal{F}f$ and the inverse Fourier transform $\mathcal{F}^{-1}f$ of $f \in \mathcal{S}(\mathbb{R}^n)$ by

$$\mathcal{F}f(\xi) = \widehat{f}(\xi) = \int_{\mathbb{R}^n} e^{-i\xi \cdot x} f(x) dx \quad \text{and} \quad \mathcal{F}^{-1}f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} f(\xi) d\xi.$$

Let $\sigma(x, \xi) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^n)$. We denote by $\mathcal{F}_1\sigma(y, \xi)$ and $\mathcal{F}_2\sigma(x, \eta)$ the partial Fourier transforms of σ in the first variable and in the second variable, respectively. That is, $\mathcal{F}_1\sigma(y, \xi) = \mathcal{F}[\sigma(\cdot, \xi)](y)$ and $\mathcal{F}_2\sigma(x, \eta) = \mathcal{F}[\sigma(x, \cdot)](\eta)$. We also denote by $\mathcal{F}_1^{-1}\sigma$ and $\mathcal{F}_2^{-1}\sigma$ the partial inverse Fourier transforms of σ in the first variable and in the second variable, respectively. We write $\mathcal{F}_{1,2} = \mathcal{F}_1\mathcal{F}_2$ and $\mathcal{F}_{1,2}^{-1} = \mathcal{F}_1^{-1}\mathcal{F}_2^{-1}$, and note that $\mathcal{F}_{1,2}$ and $\mathcal{F}_{1,2}^{-1}$ are the usual Fourier transform and inverse Fourier transform of functions on $\mathbb{R}^n \times \mathbb{R}^n$.

We introduce the α -modulation spaces based on Borup-Nielsen [1, 2]. Let $B(\xi, r)$ be the ball with center ξ and radius r , where $\xi \in \mathbb{R}^n$ and $r > 0$. A countable set \mathcal{Q} of subsets $Q \subset \mathbb{R}^n$ is called an admissible covering if $\mathbb{R}^n = \cup_{Q \in \mathcal{Q}} Q$ and there exists a constant n_0 such that $\#\{Q' \in \mathcal{Q} : Q \cap Q' \neq \emptyset\} \leq n_0$ for all $Q \in \mathcal{Q}$. We denote by $|Q|$ the Lebesgue measure of Q , and set $\langle \xi \rangle = (1 + |\xi|^2)^{1/2}$, where $\xi \in \mathbb{R}^n$. Let $0 \leq \alpha \leq 1$,

$$(2.4) \quad \begin{aligned} r_Q &= \sup\{r > 0 : B(c_r, r) \subset Q \text{ for some } c_r \in \mathbb{R}^n\}, \\ R_Q &= \inf\{R > 0 : Q \subset B(c_R, R) \text{ for some } c_R \in \mathbb{R}^n\}. \end{aligned}$$

We say that an admissible covering \mathcal{Q} is an α -covering of \mathbb{R}^n if $|Q| \asymp \langle \xi \rangle^{\alpha n}$ (uniformly) for all $\xi \in Q$ and $Q \in \mathcal{Q}$, and there exists a constant $K \geq 1$ such that $R_Q/r_Q \leq K$ for all $Q \in \mathcal{Q}$, where “ $|Q| \asymp \langle \xi \rangle^{\alpha n}$ (uniformly) for all $\xi \in Q$ and $Q \in \mathcal{Q}$ ” means that there exists a constant $C > 0$ such that

$$C^{-1} \langle \xi \rangle^{\alpha n} \leq |Q| \leq C \langle \xi \rangle^{\alpha n} \quad \text{for all } \xi \in Q \text{ and } Q \in \mathcal{Q}.$$

Let r_Q and R_Q be as in (2.4). We note that

$$(2.5) \quad B(c_Q, r_Q/2) \subset Q \subset B(d_Q, 2R_Q) \quad \text{for some } c_Q, d_Q \in \mathbb{R}^n,$$

and there exists a constant $\kappa_1 > 0$ such that

$$(2.6) \quad |Q| \geq \kappa_1 \quad \text{for all } Q \in \mathcal{Q}$$

since $|Q| \asymp \langle \xi_Q \rangle^{\alpha n} \geq 1$, where $\xi_Q \in Q$. By (2.4), we see that $s_n r_Q^n \leq |Q| \leq s_n R_Q^n$, where s_n is the volume of the unit ball in \mathbb{R}^n . This implies

$$s_n \leq \frac{|Q|}{r_Q^n} = \frac{R_Q^n}{r_Q^n} \frac{|Q|}{R_Q^n} \leq K^n \frac{|Q|}{R_Q^n} \leq K^n s_n,$$

that is,

$$(2.7) \quad |Q| \asymp r_Q^n \asymp R_Q^n \quad \text{for all } Q \in \mathcal{Q}$$

(see [1, Appendix B]). It follows from (2.6) and (2.7) that there exists a constant $\kappa_2 > 0$ such that

$$(2.8) \quad R_Q \geq \kappa_2 \quad \text{for all } Q \in \mathcal{Q}.$$

We also use the fact

$$(2.9) \quad \langle \xi_Q \rangle \asymp \langle \xi'_Q \rangle \quad \text{for all } \xi_Q, \xi'_Q \in Q \text{ and } Q \in \mathcal{Q}.$$

If $\alpha \neq 0$, then (2.9) follows directly from the definition of α -covering $|Q| \asymp \langle \xi_Q \rangle^{\alpha n}$. By (2.7), if $\alpha = 0$ then $R_Q^n \asymp |Q| \asymp \langle \xi_Q \rangle^{\alpha n} = 1$, and consequently there exists $R > 0$ such that $R_Q \leq R$ for all $Q \in \mathcal{Q}$. Hence, by (2.5), we have $Q \subset B(d_Q, 2R)$ for some $d_Q \in \mathbb{R}^n$. This implies that (2.9) is true even if $\alpha = 0$.

Given an α -covering \mathcal{Q} of \mathbb{R}^n , we say that $\{\psi_Q\}_{Q \in \mathcal{Q}}$ is a corresponding bounded admissible partition of unity (BAPU) if $\{\psi_Q\}_{Q \in \mathcal{Q}}$ satisfies

- (1) $\text{supp } \psi_Q \subset Q$,
- (2) $\sum_{Q \in \mathcal{Q}} \psi_Q(\xi) = 1$ for all $\xi \in \mathbb{R}^n$,
- (3) $\sup_{Q \in \mathcal{Q}} \|\mathcal{F}^{-1} \psi_Q\|_{L^1} < \infty$.

We remark that an α -covering \mathcal{Q} of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$ actually exists for every $0 \leq \alpha \leq 1$ ([1, Proposition A.1]). Let $1 \leq p, q \leq \infty$, $s \in \mathbb{R}$, $0 \leq \alpha \leq 1$ and \mathcal{Q} be an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$. Fix a sequence $\{\xi_Q\}_{Q \in \mathcal{Q}} \subset \mathbb{R}^n$ satisfying $\xi_Q \in Q$ for every $Q \in \mathcal{Q}$. Then the α -modulation space $M_{s,\alpha}^{p,q}(\mathbb{R}^n)$ consists of all $f \in \mathcal{S}'(\mathbb{R}^n)$ such that

$$\|f\|_{M_{s,\alpha}^{p,q}} = \left(\sum_{Q \in \mathcal{Q}} \langle \xi_Q \rangle^{sq} \|\psi_Q(D)f\|_{L^p}^q \right)^{1/q} < \infty$$

(with obvious modification in the case $q = \infty$), where $\psi(D)f = \mathcal{F}^{-1}[\psi \widehat{f}] = (\mathcal{F}^{-1}\psi) * f$. We remark that the definition of $M_{s,\alpha}^{p,q}$ is independent of the choice of the α -covering \mathcal{Q} , BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}}$ and sequence $\{\xi_Q\}_{Q \in \mathcal{Q}}$ (see [1, 2, Section 2]). Let $\psi \in \mathcal{S}(\mathbb{R}^n)$ be such that

$$(2.10) \quad \text{supp } \psi \subset [-1, 1]^n, \quad \sum_{k \in \mathbb{Z}^n} \psi(\xi - k) = 1 \quad \text{for all } \xi \in \mathbb{R}^n.$$

If $\alpha = 0$ then the α -modulation space $M_{s,\alpha}^{p,q}(\mathbb{R}^n)$ coincides with the modulation space $M_s^{p,q}(\mathbb{R}^n)$, that is, $\|f\|_{M_{s,\alpha}^{p,q}} \asymp \|f\|_{M_s^{p,q}}$, where

$$\|f\|_{M_s^{p,q}} = \left(\sum_{k \in \mathbb{Z}^n} \langle k \rangle^{sq} \|\psi(D - k)f\|_{L^p}^q \right)^{1/q}.$$

If $s = 0$, then we write $M^{p,q}(\mathbb{R}^n)$ instead of $M_0^{p,q}(\mathbb{R}^n)$. Let $\varphi_0, \varphi \in \mathcal{S}(\mathbb{R}^n)$ be such that

$$(2.11) \quad \text{supp } \varphi_0 \subset \{|\xi| \leq 2\}, \quad \text{supp } \varphi \subset \{1/2 \leq |\xi| \leq 2\}, \quad \varphi_0(\xi) + \sum_{j=1}^{\infty} \varphi(2^{-j}\xi) = 1$$

for all $\xi \in \mathbb{R}^n$, and set $\varphi_j(\xi) = \varphi(\xi/2^j)$ if $j \geq 1$. On the other hand, if $\alpha = 1$ then the α -modulation space $M_{s,\alpha}^{p,q}(\mathbb{R}^n)$ coincides with the Besov space $B_s^{p,q}(\mathbb{R}^n)$, that is, $\|f\|_{M_{s,\alpha}^{p,q}} \asymp \|f\|_{B_s^{p,q}}$, where

$$\|f\|_{B_s^{p,q}} = \left(\sum_{j=0}^{\infty} 2^{jsq} \|\varphi_j(D)f\|_{L^p}^q \right)^{1/q}.$$

We remark that we can actually check that the α -covering \mathcal{Q} with the corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$ given in [1, Proposition A.1] satisfies

$$(2.12) \quad \sum_{Q \in \mathcal{Q}} \psi_Q(D)f = f \quad \text{in } \mathcal{S}'(\mathbb{R}^n) \quad \text{for all } f \in \mathcal{S}'(\mathbb{R}^n)$$

and

$$(2.13) \quad \sum_{Q \in \mathcal{Q}} \sum_{Q' \in \mathcal{Q}} \psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma(x, \xi) = \sigma(x, \xi) \quad \text{in } \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$$

for all $\sigma \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$, where $0 \leq \alpha < 1$,

$$\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma = \mathcal{F}_{1,2}^{-1}[(\psi_Q \otimes \psi_{Q'}) \mathcal{F}_{1,2} \sigma] = [(\mathcal{F}^{-1} \psi_Q) \otimes (\mathcal{F}^{-1} \psi_{Q'})] * \sigma$$

and $\psi_Q \otimes \psi_{Q'}(x, \xi) = \psi_Q(x) \psi_{Q'}(\xi)$. In the case $\alpha = 1$, (2.12) and (2.13) are well known facts, since we can take $\{\varphi_j\}_{j \geq 0}$ as a BAPU corresponding to the α -covering $\{\{|\xi| \leq 2\}, \{2^{j-1} \leq |\xi| \leq 2^{j+1}\}\}_{j \geq 1}\}$, where $\{\varphi_j\}_{j \geq 0}$ is as in (2.11). In the rest of this paper, we assume that an α -covering \mathcal{Q} with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$ always satisfies (2.12) and (2.13).

We introduce the product α -modulation spaces $M_{(s_1, s_2), (\alpha, \alpha)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$ as symbol classes of pseudo-differential operators. Let $1 \leq p, q \leq \infty$, $s_1, s_2 \in \mathbb{R}$, $0 \leq \alpha \leq 1$ and \mathcal{Q} be an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$. Fix two sequences $\{x_Q\}_{Q \in \mathcal{Q}}, \{\xi_{Q'}\}_{Q' \in \mathcal{Q}} \subset \mathbb{R}^n$ satisfying $x_Q \in Q$ and $\xi_{Q'} \in Q'$ for every $Q, Q' \in \mathcal{Q}$. Then the product α -modulation space $M_{(s_1, s_2), (\alpha, \alpha)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$ consists of all $\sigma \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$ such that

$$\|\sigma\|_{M_{(s_1, s_2), (\alpha, \alpha)}^{(p, p), (q, q)}} = \left\{ \sum_{Q \in \mathcal{Q}} \sum_{Q' \in \mathcal{Q}} (\langle x_Q \rangle^{s_1} \langle \xi_{Q'} \rangle^{s_2} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma\|_{L^p(\mathbb{R}^n \times \mathbb{R}^n)})^q \right\}^{1/q} < \infty$$

(with obvious modification in the case $q = \infty$). Since we can take $\{\psi(\cdot - k)\}_{k \in \mathbb{Z}^n}$ as a BAPU corresponding to the α -covering $\{k + [-1, 1]^n\}_{k \in \mathbb{Z}^n}$ if $\alpha = 0$, we have $M_{(s_1, s_2), (0, 0)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n) = M_{(s_1, s_2)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$, where

$$\|\sigma\|_{M_{(s_1, s_2)}^{(p, p), (q, q)}} = \left\{ \sum_{k \in \mathbb{Z}^n} \sum_{\ell \in \mathbb{Z}^n} (\langle k \rangle^{s_1} \langle \ell \rangle^{s_2} \|\psi(D_x - k) \psi(D_\xi - \ell) \sigma\|_{L^p(\mathbb{R}^n \times \mathbb{R}^n)})^q \right\}^{1/q}$$

and $\psi \in \mathcal{S}(\mathbb{R}^n)$ is as in (2.10). In particular, the space $M_{(0, 0), (0, 0)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$ of product type on $\mathbb{R}^n \times \mathbb{R}^n$ coincides with the ordinary modulation space $M^{p, q}(\mathbb{R}^{2n})$ on \mathbb{R}^{2n} . Here we have used the fact that $\psi \otimes \psi$ satisfies (2.10) with $2n$ instead of n . Similarly, $M_{(s_1, s_2), (1, 1)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n) = B_{(s_1, s_2)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$, where

$$\|\sigma\|_{B_{(s_1, s_2)}^{(p, p), (q, q)}} = \left\{ \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} (2^{js_1 + ks_2} \|\varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^p(\mathbb{R}^n \times \mathbb{R}^n)})^q \right\}^{1/q}$$

and $\{\varphi_j\}_{j \geq 0}$ is as in (2.11) (see Sugimoto [25, p.116]). Hereafter, we simply write $M_{(s_1, s_2), \alpha}^{p, q}(\mathbb{R}^n \times \mathbb{R}^n)$ instead of $M_{(s_1, s_2), (\alpha, \alpha)}^{(p, p), (q, q)}(\mathbb{R}^n \times \mathbb{R}^n)$, where $\mathbf{p} = (p, p)$, $\mathbf{q} = (q, q)$ and $\boldsymbol{\alpha} = (\alpha, \alpha)$.

We remark the following basic facts, and give the proof in Appendix B for reader's convenience.

Lemma 2.1 ([17, Lemma 2.1]). *Let \mathcal{Q} be an α -covering of \mathbb{R}^n and $R > 0$. Then the following are true:*

- (1) *If $(Q + B(0, R)) \cap Q' \neq \emptyset$, then there exists a constant $\kappa > 0$ such that*

$$\kappa^{-1} \langle \xi_Q \rangle \leq \langle \xi_{Q, Q'} \rangle \leq \kappa \langle \xi_Q \rangle \quad \text{and} \quad \kappa^{-1} \langle \xi_{Q'} \rangle \leq \langle \xi_{Q, Q'} \rangle \leq \kappa \langle \xi_{Q'} \rangle$$

for all $\xi_Q \in Q$, $\xi_{Q'} \in Q'$ and $\xi_{Q, Q'} \in (Q + B(0, R)) \cap Q'$, where κ is independent of $Q, Q' \in \mathcal{Q}$. In particular, $\langle \xi_Q \rangle \asymp \langle \xi_{Q'} \rangle$.

- (2) *There exists a constant n'_0 such that*

$$\#\{Q' \in \mathcal{Q} : (Q + B(0, R)) \cap Q' \neq \emptyset\} \leq n'_0 \quad \text{for all } Q \in \mathcal{Q}.$$

3. TRACE PROPERTY OF PSEUDO-DIFFERENTIAL OPERATORS

In this section, we prove Theorem 1.1. For $\sigma \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$, the pseudo-differential operator $\sigma(X, D)$ is defined by

$$\sigma(X, D)f(x) = \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \sigma(x, \xi) \widehat{f}(\xi) d\xi \quad \text{for } f \in \mathcal{S}(\mathbb{R}^n).$$

We define the Rihaczek distribution $R(f, g)$ of f and g by

$$R(f, g)(x, \xi) = f(x) \overline{\widehat{g}(\xi)} e^{-ix \cdot \xi} \quad \text{for } x, \xi \in \mathbb{R}^n.$$

Then

$$\langle \sigma(X, D)f, g \rangle = (2\pi)^{-n} \langle \sigma, R(g, f) \rangle \quad \text{for all } f, g \in \mathcal{S}(\mathbb{R}^n).$$

Gröchenig proved that $\sigma(X, D)$ is a trace operator if $\sigma \in M^{1,1}(\mathbb{R}^{2n})$, and the Rihaczek distribution plays an important role in his proof [13].

Let $0 \leq \alpha \leq 1$ and \mathcal{Q} be an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$. In order to prove Theorem 1.1, we introduce a modified version of Rihaczek distribution $R_{Q, Q'}(f, g)$ of f and g defined by

$$R_{Q, Q'}(f, g)(x, \xi) = f(x) \overline{\widehat{g}(\xi)} e^{-i(x/R_Q) \cdot (\xi/R_{Q'})} \quad \text{for } x, \xi \in \mathbb{R}^n,$$

where $f, g \in \mathcal{S}(\mathbb{R}^n)$, $Q, Q' \in \mathcal{Q}$, and $R_Q, R_{Q'}$ are as in (2.4). We denote by $\widehat{R}_{Q, Q'}(f, g)$ the Fourier transform of $R_{Q, Q'}(f, g)$ in both variables $x, \xi \in \mathbb{R}^n$, that is, $\widehat{R}_{Q, Q'}(f, g) = \mathcal{F}_{1,2} R_{Q, Q'}(f, g)$.

Lemma 3.1. *Let $f, g \in \mathcal{S}(\mathbb{R}^n)$. Then*

$$\widehat{R}_{Q, Q'}(f, g)(y, \eta) = \int_{\mathbb{R}^n} e^{-i\eta \cdot \xi} \widehat{f}(y + (\xi/R_Q R_{Q'})) \overline{\widehat{g}(\xi)} d\xi.$$

Proof. By Fubini's theorem,

$$\begin{aligned} \widehat{R}_{Q, Q'}(f, g)(y, \eta) &= \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} e^{-i(y \cdot x + \eta \cdot \xi)} R_{Q, Q'}(f, g)(x, \xi) dx d\xi \\ &= \int_{\mathbb{R}^n} e^{-i\eta \cdot \xi} \overline{\widehat{g}(\xi)} \left(\int_{\mathbb{R}^n} e^{-i(y + (\xi/R_Q R_{Q'})) \cdot x} f(x) dx \right) d\xi \\ &= \int_{\mathbb{R}^n} e^{-i\eta \cdot \xi} \widehat{f}(y + (\xi/R_Q R_{Q'})) \overline{\widehat{g}(\xi)} d\xi. \end{aligned}$$

The proof is complete. \square

Let $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^n) \setminus \{0\}$ be such that

$$(3.1) \quad \widehat{\varphi}_1, \widehat{\varphi}_2 \geq 0, \quad \widehat{\varphi}_1 \geq 1 \text{ on } \{\xi : |\xi| \leq 4 + 1/4\kappa_2^2\}, \quad \text{supp } \widehat{\varphi}_2 \subset \{\xi : |\xi| \leq 1/4\},$$

where κ_2 is as in (2.8).

Lemma 3.2. *Let $\varphi_1, \varphi_2 \in \mathcal{S}(\mathbb{R}^n)$ be as in (3.1). Then the following are true:*

- (1) *For every $\alpha, \beta \in \mathbb{Z}_+^n$, $\sup_{Q, Q' \in \mathcal{Q}} \|\partial_y^\alpha \partial_\eta^\beta \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)\|_{L^\infty(\mathbb{R}^n \times \mathbb{R}^n)} < \infty$.*
- (2) *There exists a constant $C > 0$ such that $|\widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta)| \geq C$ for all $Q, Q' \in \mathcal{Q}$, $|y| \leq 4$ and $|\eta| \leq 4$.*

Proof. By Lemma 3.1,

$$\partial_y^\alpha \partial_\eta^\beta \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta) = \int_{\mathbb{R}^n} e^{-i\eta \cdot \xi} (-i\xi)^\beta (\partial^\alpha \widehat{\varphi}_1)(y + (\xi/R_Q R_{Q'})) \overline{\widehat{\varphi}_2(\xi)} d\xi.$$

Hence,

$$|\partial_y^\alpha \partial_\eta^\beta \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta)| \leq 4^{-|\beta|} \|\partial^\alpha \widehat{\varphi}_1\|_{L^\infty} \|\widehat{\varphi}_2\|_{L^1}$$

for all $y, \eta \in \mathbb{R}^n$ and $Q, Q' \in \mathcal{Q}$, and this is the first part.

We next consider the second part. Note that $\cos(\eta \cdot \xi) \geq C > 0$ for all $|\eta| \leq 4$ and $|\xi| \leq 1/4$ since $|\eta \cdot \xi| \leq 1$. Similarly, $\widehat{\varphi}_1(y + (\xi/R_Q R_{Q'})) \geq 1$ for all $|y| \leq 4$ and $|\xi| \leq 1/4$ since $|y + (\xi/R_Q R_{Q'})| \leq 4 + 1/4\kappa_2^2$, where κ_2 is as in (3.1). Therefore, by Lemma 3.1 and our assumption $\widehat{\varphi}_1, \widehat{\varphi}_2 \geq 0$, we have

$$\begin{aligned} & |\widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta)| \\ &= \left| \int_{\mathbb{R}^n} (\cos(\eta \cdot \xi) - i \sin(\eta \cdot \xi)) \widehat{\varphi}_1(y + (\xi/R_Q R_{Q'})) \overline{\widehat{\varphi}_2(\xi)} d\xi \right| \\ &\geq \left| \int_{\mathbb{R}^n} \cos(\eta \cdot \xi) \widehat{\varphi}_1(y + (\xi/R_Q R_{Q'})) \widehat{\varphi}_2(\xi) d\xi \right| \\ &= \left| \int_{|\xi| \leq 1/4} \cos(\eta \cdot \xi) \widehat{\varphi}_1(y + (\xi/R_Q R_{Q'})) \widehat{\varphi}_2(\xi) d\xi \right| \\ &\geq C \int_{|\xi| \leq 1/4} \widehat{\varphi}_1(y + (\xi/R_Q R_{Q'})) \widehat{\varphi}_2(\xi) d\xi \geq C \int_{|\xi| \leq 1/4} \widehat{\varphi}_2(\xi) d\xi = C \|\widehat{\varphi}_2\|_{L^1} \end{aligned}$$

for all $|y|, |\eta| \leq 4$. The proof is complete. \square

Let φ_1, φ_2 be as in (3.1), and set

$$(3.2) \quad \varphi_{Q, Q'}(y, \eta) = \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)((y - d_Q)/R_Q, (\eta - d_{Q'})/R_{Q'}),$$

where $d_Q, d_{Q'}, R_Q, R_{Q'}$ are as in (2.5). We denote by T_x and M_ξ the operators of translation and modulation:

$$T_x f(t) = f(t - x), \quad M_\xi f(t) = e^{i\xi \cdot t} f(t),$$

where $x, \xi, t \in \mathbb{R}^n$.

Lemma 3.3. *Let $\Phi_{Q, Q'} = \mathcal{F}_{1,2}^{-1} \varphi_{Q, Q'}$, where $\varphi_{Q, Q'}$ is defined by (3.2). Then*

$$\Phi_{Q, Q'}(x, \xi) = R_Q^n (M_{d_Q/R_Q} \varphi_1)(R_Q x) \overline{\mathcal{F}[(T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})](\xi)} e^{-ix \cdot \xi}$$

for all $Q, Q' \in \mathcal{Q}$.

Proof. A straightforward computation shows that

$$\begin{aligned}
\Phi_{Q,Q'}(x, \xi) &= \mathcal{F}_{(y,\eta) \rightarrow (x,\xi)}^{-1} \left[\widehat{R}_{Q,Q'}(\varphi_1, \varphi_2)((y - d_Q)/R_Q, (\eta - d_{Q'})/R_{Q'}) \right] \\
&= R_Q^n R_{Q'}^n e^{i(d_Q \cdot x + d_{Q'} \cdot \xi)} \mathcal{F}_{1,2}^{-1} \widehat{R}_{Q,Q'}(\varphi_1, \varphi_2)(R_Q x, R_{Q'} \xi) \\
&= R_Q^n R_{Q'}^n e^{i(d_Q \cdot x + d_{Q'} \cdot \xi)} R_{Q,Q'}(\varphi_1, \varphi_2)(R_Q x, R_{Q'} \xi) \\
&= R_Q^n R_{Q'}^n e^{i(d_Q \cdot x + d_{Q'} \cdot \xi)} \varphi_1(R_Q x) \overline{\varphi_2(R_{Q'} \xi)} e^{-i(R_Q x/R_Q) \cdot (R_{Q'} \xi/R_{Q'})} \\
&= R_Q^n (M_{d_Q/R_Q} \varphi_1)(R_Q x) \overline{R_{Q'}^n \mathcal{F}[T_{d_{Q'}/R_{Q'}} \varphi_2](R_{Q'} \xi)} e^{-ix \cdot \xi}.
\end{aligned}$$

This completes the proof. \square

Lemma 3.4. *Let $\Phi_{Q,Q'} = \mathcal{F}_{1,2}^{-1} \varphi_{Q,Q'}$, where $\varphi_{Q,Q'}$ is defined by (3.2). Then there exists a constant $C > 0$ such that*

$$\|\Phi_{Q,Q'}(X - y, D - \eta)\|_{\mathcal{L}_1} \leq C|Q|^{1/2}|Q'|^{1/2}$$

for all $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$.

Proof. By Lemma 3.3,

$$\begin{aligned}
\Phi_{Q,Q'}(x - y, \xi - \eta) &= R_Q^n (M_{d_Q/R_Q} \varphi_1)(R_Q(x - y)) \\
&\quad \times \overline{\mathcal{F}[(T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})](\xi - \eta)} e^{-i(x-y) \cdot (\xi - \eta)} \\
&= R_Q^n e^{iy \cdot \eta} (T_{R_Q y} M_{d_Q/R_Q} \varphi_1)(R_Q x) \\
&\quad \times \overline{e^{-iy \cdot \xi} \mathcal{F}[(M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})](\xi)} e^{-i(x \cdot \xi + y \cdot \eta)} \\
&= R_Q^n (M_{\eta/R_Q} T_{R_Q y} M_{d_Q/R_Q} \varphi_1)(R_Q x) \\
&\quad \times \overline{\mathcal{F}[(T_{y/R_Q} M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})](\xi)} e^{-i(x \cdot \xi + y \cdot \eta)}.
\end{aligned}$$

Hence,

$$\begin{aligned}
&\Phi_{Q,Q'}(X - y, D - \eta) f(x) \\
&= e^{-iy \cdot \eta} (2\pi)^{-n} \langle \widehat{f}, \mathcal{F}[(T_{y/R_Q} M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})] \rangle \\
&\quad \times R_Q^n (M_{\eta/R_Q} T_{R_Q y} M_{d_Q/R_Q} \varphi_1)(R_Q x) \\
&= e^{-iy \cdot \eta} \langle f, (T_{y/R_Q} M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'}) \rangle \\
&\quad \times R_Q^n (M_{\eta/R_Q} T_{R_Q y} M_{d_Q/R_Q} \varphi_1)(R_Q x),
\end{aligned}$$

and consequently $\Phi_{Q,Q'}(X - y, D - \eta)$ is a rank one operator. By (2.7) and Schwarz's inequality, we have

$$\begin{aligned}
\|\Phi_{Q,Q'}(X - y, D - \eta) f\|_{L^2} &\leq R_Q^n \|(M_{\eta/R_Q} T_{R_Q y} M_{d_Q/R_Q} \varphi_1)(R_Q \cdot)\|_{L^2} \\
&\quad \times \|(T_{y/R_Q} M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2)(\cdot/R_{Q'})\|_{L^2} \|f\|_{L^2} \\
&= R_Q^{n/2} R_{Q'}^{n/2} \|M_{\eta/R_Q} T_{R_Q y} M_{d_Q/R_Q} \varphi_1\|_{L^2} \\
&\quad \times \|T_{y/R_Q} M_{R_{Q'} \eta} T_{d_{Q'}/R_{Q'}} \varphi_2\|_{L^2} \|f\|_{L^2} \\
&= R_Q^{n/2} R_{Q'}^{n/2} \|\varphi_1\|_{L^2} \|\varphi_2\|_{L^2} \|f\|_{L^2} \\
&\leq C|Q|^{1/2}|Q'|^{1/2} \|f\|_{L^2}
\end{aligned}$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$, $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$. Therefore,

$$\|\Phi_{Q,Q'}(X - y, D - \eta)\|_{\mathcal{L}_1} = \|\Phi_{Q,Q'}(X - y, D - \eta)\|_{\mathcal{L}(L^2)} \leq C|Q|^{1/2}|Q'|^{1/2}$$

for all $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$. The proof is complete. \square

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. By (2.13),

$$(3.3) \quad \|\sigma(X, D)\|_{\mathcal{I}_1} \leq \sum_{Q, Q' \in \mathcal{Q}} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma(X, D)\|_{\mathcal{I}_1},$$

where \mathcal{Q} is an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$. Let $\gamma \in \mathcal{S}(\mathbb{R}^n)$ be such that $\gamma = 1$ on $\{\xi : |\xi| \leq 2\}$ and $\text{supp } \gamma \subset \{\xi : |\xi| \leq 4\}$, and set

$$\gamma_{Q, Q'}(y, \eta) = \gamma((y - d_Q)/R_Q) \gamma((\eta - d_{Q'})/R_{Q'}),$$

where $d_Q, d_{Q'}, R_Q, R_{Q'}$ are as in (2.5). Recall that $\text{supp } \psi_Q \subset Q$ for all $Q \in \mathcal{Q}$ (see the definition of BAPU). Since $\gamma_{Q, Q'}(y, \eta) = 1$ on $\{(y, \eta) : |y - d_Q| \leq 2R_Q, |\eta - d_{Q'}| \leq 2R_{Q'}\}$, we have by (2.5)

$$(3.4) \quad \psi_Q(y) \psi_{Q'}(\eta) = \gamma_{Q, Q'}(y, \eta) \psi_Q(y) \psi_{Q'}(\eta)$$

for all $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$. On the other hand, since $\text{supp } \gamma \otimes \gamma \subset \{(y, \eta) : |y| \leq 4, |\eta| \leq 4\}$, we have by Lemma 3.2 (2)

$$\gamma(y) \gamma(\eta) = \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta) \frac{\gamma(y) \gamma(\eta)}{\widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta)}$$

for all $y, \eta \in \mathbb{R}^n$, where φ_1, φ_2 are as in (3.1). This implies

$$(3.5) \quad \begin{aligned} \gamma_{Q, Q'}(y, \eta) &= \gamma((y - d_Q)/R_Q) \gamma((\eta - d_{Q'})/R_{Q'}) \\ &= \widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)((y - d_Q)/R_Q, (\eta - d_{Q'})/R_{Q'}) \\ &\quad \times \frac{\gamma((y - d_Q)/R_Q) \gamma((\eta - d_{Q'})/R_{Q'})}{\widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)((y - d_Q)/R_Q, (\eta - d_{Q'})/R_{Q'})} \\ &= \varphi_{Q, Q'}(y, \eta) \frac{\gamma_{Q, Q'}(y, \eta)}{\varphi_{Q, Q'}(y, \eta)} \end{aligned}$$

for all $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$, where $\varphi_{Q, Q'}$ is defined by (3.2). Combining (3.4) and (3.5), we see that

$$\psi_Q(y) \psi_{Q'}(\eta) = \varphi_{Q, Q'}(y, \eta) \frac{\gamma_{Q, Q'}(y, \eta)}{\varphi_{Q, Q'}(y, \eta)} \psi_Q(y) \psi_{Q'}(\eta)$$

for all $Q, Q' \in \mathcal{Q}$ and $y, \eta \in \mathbb{R}^n$. Then

$$(3.6) \quad \begin{aligned} &\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma(x, \xi) \\ &= \int_{\mathbb{R}^{2n}} \Phi_{Q, Q'}(x - y, \xi - \eta) \left[\mathcal{F}_{1,2}^{-1} \left(\frac{\gamma_{Q, Q'}}{\varphi_{Q, Q'}} \right) \right] * [\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma](y, \eta) dy d\eta, \end{aligned}$$

where $\Phi_{Q, Q'} = \mathcal{F}_{1,2}^{-1} \varphi_{Q, Q'}$. We note that

$$(3.7) \quad \sup_{Q, Q' \in \mathcal{Q}} \left\| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma_{Q, Q'}}{\varphi_{Q, Q'}} \right) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} < \infty.$$

In fact, by Lemma 3.2,

$$\sup_{y, \eta \in \mathbb{R}^n} \left| \partial_y^\alpha \partial_\eta^\beta \left(\frac{\gamma(y) \gamma(\eta)}{\widehat{R}_{Q, Q'}(\varphi_1, \varphi_2)(y, \eta)} \right) \right| \leq C_{\alpha, \beta} \quad \text{for all } Q, Q' \in \mathcal{Q},$$

where $|\alpha + \beta| \leq 2n + 1$. Hence, using $\text{supp } \gamma \otimes \gamma / \widehat{R}_{Q,Q'}(\varphi_1, \varphi_2) \subset \{(y, \eta) : |y| \leq 4, |\eta| \leq 4\}$ and integration by parts, we have

$$(3.8) \quad \sup_{Q, Q' \in \mathcal{Q}} \left\| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma \otimes \gamma}{\widehat{R}_{Q,Q'}(\varphi_1, \varphi_2)} \right) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} < \infty.$$

On the other hand, by a change of variables, we see that

$$(3.9) \quad \begin{aligned} & \left\| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma_{Q,Q'}}{\varphi_{Q,Q'}} \right) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &= R_Q^n R_{Q'}^n \left\| \left[\mathcal{F}_{1,2}^{-1} \left(\frac{\gamma \otimes \gamma}{\widehat{R}_{Q,Q'}(\varphi_1, \varphi_2)} \right) \right] (R_Q x, R_{Q'} \xi) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &= \left\| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma \otimes \gamma}{\widehat{R}_{Q,Q'}(\varphi_1, \varphi_2)} \right) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)}. \end{aligned}$$

Combining (3.8) and (3.9), we obtain (3.7). Recall that $\langle x_Q \rangle^{\alpha n} \asymp |Q|$ and $\langle \xi_{Q'} \rangle^{\alpha n} \asymp |Q'|$ for all $Q, Q' \in \mathcal{Q}$, where $x_Q \in Q$ and $\xi_{Q'} \in Q'$ (see the definition of an α -covering). By (3.6), (3.7) and Lemma 3.4, we see that

$$\begin{aligned} & \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma(X, D)\|_{\mathcal{I}_1} \\ & \leq \int_{\mathbb{R}^{2n}} \|\Phi_{Q,Q'}(X - y, D - \eta)\|_{\mathcal{I}_1} \left| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma_{Q,Q'}}{\varphi_{Q,Q'}} \right) * [\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma](y, \eta) \right| dy d\eta \\ & \leq C |Q|^{1/2} |Q'|^{1/2} \left\| \mathcal{F}_{1,2}^{-1} \left(\frac{\gamma_{Q,Q'}}{\varphi_{Q,Q'}} \right) \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ & \leq C \langle x_Q \rangle^{\alpha n/2} \langle \xi_{Q'} \rangle^{\alpha n/2} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \end{aligned}$$

for all $Q, Q' \in \mathcal{Q}$. Therefore, by (3.3), we have

$$\|\sigma(X, D)\|_{\mathcal{I}_1} \leq C \sum_{Q, Q' \in \mathcal{Q}} \langle x_Q \rangle^{\alpha n/2} \langle \xi_{Q'} \rangle^{\alpha n/2} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)},$$

where C is independent of σ . The proof is complete.

4. TRACE PROPERTY OF COMMUTATORS

In this section, we prove Theorem 1.2. We recall the definition of commutators. Let a be a Lipschitz function on \mathbb{R}^n , that is,

$$(4.1) \quad |a(x) - a(y)| \leq A|x - y| \quad \text{for all } x, y \in \mathbb{R}^n.$$

Note that a satisfies (4.1) if and only if a is differentiable (in the ordinary sense) and $\partial^\beta a \in L^\infty(\mathbb{R}^n)$ for $|\beta| = 1$ (see [24, Chapter 8, Theorem 3]). If T is a bounded linear operator on $L^2(\mathbb{R}^n)$, then $T(af)$ and $a(Tf)$ make sense as elements in $L^2_{\text{loc}}(\mathbb{R}^n)$ when $f \in \mathcal{S}(\mathbb{R}^n)$, since $|a(x)| \leq C(1 + |x|)$ for some constant $C > 0$. Hence, the commutator $[T, a]$ can be defined by

$$[T, a]f(x) = T(af)(x) - a(x)Tf(x) \quad \text{for } f \in \mathcal{S}(\mathbb{R}^n),$$

where T is a bounded linear operator on $L^2(\mathbb{R}^n)$. In order to prove Theorem 1.2, we prepare the following lemmas:

Lemma 4.1 ([17, Lemma 4.1]). *Let T be a bounded linear operator on $L^2(\mathbb{R}^n)$, and a be a Lipschitz function on \mathbb{R}^n with $\|\nabla a\|_{L^\infty} \neq 0$. Then there exist $\epsilon(a) > 0$ and $\{a_\epsilon\}_{0 < \epsilon < \epsilon(a)} \subset \mathcal{S}(\mathbb{R}^n)$ such that*

- (1) $\langle [T, a]f, g \rangle = \lim_{\epsilon \rightarrow 0} \langle [T, a_\epsilon]f, g \rangle$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$,
- (2) $\|\nabla a_\epsilon\|_{L^\infty} \leq C\|\nabla a\|_{L^\infty}$ for all $0 < \epsilon < \epsilon(a)$,

where $\nabla a = (\partial_1 a, \dots, \partial_n a)$, and C is independent of T and a .

We give the proof of Lemma 4.1 in Appendix B for reader's convenience.

Lemma 4.2. *Let $\sigma(x, \xi) \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$ and $\gamma \in \mathcal{S}(\mathbb{R}^n)$ be such that $\text{supp } \widehat{\sigma}_x \subset B(\zeta, R)$ for all $x \in \mathbb{R}^n$ and $\text{supp } \widehat{\gamma} \subset B(0, 1)$, where $\sigma_x(\xi) = \sigma(x, \xi)$, $\widehat{\sigma}_x(\eta) = \mathcal{F}_2 \sigma(x, \eta)$, $\zeta \in \mathbb{R}^n$ and $R > 0$. Then there exists a constant $C > 0$ such that*

$$\int_{\mathbb{R}^{2n}} \left| \int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \xi + t\eta) \gamma(\eta) \widehat{f}(\eta) d\eta \right| dx d\xi \leq C(1 + R)^{n/2} \|\sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{L^\infty}$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$ and $0 < t < 1$, where C is independent of σ , $\zeta \in \mathbb{R}^n$ and $R > 0$.

Proof. Since $\text{supp } \mathcal{F}_{\eta \rightarrow \eta'} [\sigma(x, \xi + t\eta)] \subset tB(\zeta, R)$ and $\text{supp } \widehat{\gamma} \subset B(0, 1)$, we have

$$\text{supp } \mathcal{F}_{\eta \rightarrow \eta'} [\sigma(x, \xi + t\eta) \gamma(\eta)] \subset tB(\zeta, R) + B(0, 1) = B(t\zeta, 1 + tR)$$

for all $x, \xi \in \mathbb{R}^n$ and $0 < t < 1$, where $tB(\zeta, R) = \{t\eta' : \eta' \in B(\zeta, R)\}$. Hence, by Plancherel's theorem,

$$\begin{aligned} & \int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \xi + t\eta) \gamma(\eta) \widehat{f}(\eta) d\eta \\ &= \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} e^{-iy \cdot \eta} \sigma(x, \xi + t\eta) \gamma(\eta) d\eta \right) f(x + y) dy \\ &= \int_{\mathbb{R}^n} \mathcal{F}_{\eta \rightarrow y} [\sigma(x, \xi + t\eta) \gamma(\eta)] \chi_{B(t\zeta, 1+tR)}(y) (T_{-x} f)(y) dy \\ &= (2\pi)^n \int_{\mathbb{R}^n} \sigma(x, \xi + t\eta) \gamma(\eta) \mathcal{F}^{-1} [\chi_{B(t\zeta, 1+tR)} (T_{-x} f)](-\eta) d\eta \end{aligned}$$

for all $x, \xi \in \mathbb{R}^n$ and $0 < t < 1$, where $\chi_{B(t\zeta, 1+tR)}$ is the characteristic function of $B(t\zeta, 1+tR)$. Therefore, by Fubini's theorem, Schwarz's inequality and Plancherel's theorem, we have

$$\begin{aligned} & \int_{\mathbb{R}^{2n}} \left| \int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \xi + t\eta) \gamma(\eta) \widehat{f}(\eta) d\eta \right| dx d\xi \\ &\leq (2\pi)^n \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |\sigma(x, \xi + t\eta)| d\xi \right) |\gamma(\eta) \mathcal{F}^{-1} [\chi_{B(t\zeta, 1+tR)} (T_{-x} f)](\eta)| d\eta dx \\ &= (2\pi)^n \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^n} |\sigma(x, \xi)| d\xi \right) |\gamma(\eta) \mathcal{F}^{-1} [\chi_{B(t\zeta, 1+tR)} (T_{-x} f)](\eta)| d\eta dx \\ &= (2\pi)^n \int_{\mathbb{R}^n} \int_{\mathbb{R}^n} |\sigma(x, \xi)| \left(\int_{\mathbb{R}^n} |\gamma(\eta) \mathcal{F}^{-1} [\chi_{B(t\zeta, 1+tR)} (T_{-x} f)](\eta)| d\eta \right) dx d\xi \\ &\leq (2\pi)^n \int_{\mathbb{R}^{2n}} |\sigma(x, \xi)| (\|\gamma\|_{L^2} \|\mathcal{F}^{-1} [\chi_{B(t\zeta, 1+tR)} (T_{-x} f)]\|_{L^2}) dx d\xi \\ &= (2\pi)^{n/2} \int_{\mathbb{R}^{2n}} |\sigma(x, \xi)| (\|\gamma\|_{L^2} \|\chi_{B(t\zeta, 1+tR)} (T_{-x} f)\|_{L^2}) dx d\xi \\ &\leq (2\pi)^{n/2} \|\gamma\|_{L^2} |B(t\zeta, 1 + tR)|^{1/2} \|\sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{L^\infty} \end{aligned}$$

$$= C(1 + tR)^{n/2} \|\sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{L^\infty} \leq C(1 + R)^{n/2} \|\sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|f\|_{L^\infty}$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$ and $0 < t < 1$. The proof is complete. \square

Lemma 4.3. *Let $0 \leq \alpha \leq 1$ and \mathcal{Q} be an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$. Then, for every $\beta \in \mathbb{Z}_+^n$ there exists a constant $C_\beta > 0$ such that*

$$\|\partial^\beta(\mathcal{F}^{-1}\psi_Q)\|_{L^1} \leq C_\beta \langle \xi_Q \rangle^{|\beta|} \quad \text{for all } \xi_Q \in Q \text{ and } Q \in \mathcal{Q}.$$

Proof. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be such that $\varphi = 1$ on $B(0, 2)$, and set $\varphi_Q(\xi) = \varphi((\xi - d_Q)/R_Q)$, where $Q \in \mathcal{Q}$ and d_Q, R_Q are as in (2.5). Since $\varphi_Q = 1$ on $B(d_Q, 2R_Q)$ and $\text{supp } \psi_Q \subset Q \subset B(d_Q, 2R_Q)$, we see that

$$\mathcal{F}^{-1}\psi_Q(x) = \mathcal{F}^{-1}[\varphi_Q \psi_Q](x) = \int_{\mathbb{R}^n} e^{id_Q \cdot (x-y)} R_Q^n \Phi(R_Q(x-y)) (\mathcal{F}^{-1}\psi_Q)(y) dy$$

for all $Q \in \mathcal{Q}$, where $\Phi = \mathcal{F}^{-1}\varphi$. Hence,

$$\begin{aligned} & \partial^\beta \mathcal{F}^{-1}\psi_Q(x) \\ &= \sum_{\beta_1 + \beta_2 = \beta} C_{\beta_1, \beta_2} \int_{\mathbb{R}^n} d_Q^{\beta_1} e^{id_Q \cdot (x-y)} R_Q^{n+|\beta_2|} (\partial^{\beta_2} \Phi)(R_Q(x-y)) (\mathcal{F}^{-1}\psi_Q)(y) dy \end{aligned}$$

for all $Q \in \mathcal{Q}$. Since $R_Q \asymp |Q|^{1/n} \asymp \langle \xi_Q \rangle^\alpha$ (see (2.7)) and $\xi_Q \in B(d_Q, 2R_Q)$,

$$|d_Q| \leq |d_Q - \xi_Q| + |\xi_Q| \leq 2R_Q + \langle \xi_Q \rangle \leq C \langle \xi_Q \rangle,$$

and consequently $|d_Q| \leq C \langle \xi_Q \rangle$ for all $\xi_Q \in Q$ and $Q \in \mathcal{Q}$. Therefore,

$$\|\partial^\beta(\mathcal{F}^{-1}\psi_Q)\|_{L^1} \leq C \langle \xi_Q \rangle^{|\beta|} \left(\sum_{\beta' \leq \beta} \|\partial^{\beta'} \Phi\|_{L^1} \right) \sup_{Q \in \mathcal{Q}} \|\mathcal{F}^{-1}\psi_Q\|_{L^1} = C_\beta \langle \xi_Q \rangle^{|\beta|}$$

for all $Q \in \mathcal{Q}$. The proof is complete. \square

We are now ready to prove Theorem 1.2.

Proof of Theorem 1.2. Let $\sigma \in M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n)$. Then, by Theorem 1.1 and (2.1), we see that $\sigma(X, D)$ is bounded on $L^2(\mathbb{R}^n)$. Note that $\sigma(x, \xi) \in L^1(\mathbb{R}^n \times \mathbb{R}^n)$ since

$$\|\sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \leq \sum_{Q, Q' \in \mathcal{Q}} \|\psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \leq \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}},$$

where \mathcal{Q} is an α -covering of \mathbb{R}^n with a corresponding BAPU $\{\psi_Q\}_{Q \in \mathcal{Q}} \subset \mathcal{S}(\mathbb{R}^n)$.

We first consider the case $a \in \mathcal{S}(\mathbb{R}^n)$. Using

$$\begin{aligned} \sigma(X, D)(af)(x) &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \eta) \widehat{af}(\eta) d\eta \\ &= \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \eta) \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} \widehat{a}(\eta - \xi) \widehat{f}(\xi) d\xi \right) d\eta \\ &= \frac{1}{(2\pi)^{2n}} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \left(\int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \xi + \eta) \widehat{a}(\eta) d\eta \right) \widehat{f}(\xi) d\xi \end{aligned}$$

and

$$a(x) \sigma(X, D)f(x) = \left(\frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \eta} \widehat{a}(\eta) d\eta \right) \frac{1}{(2\pi)^n} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \sigma(x, \xi) \widehat{f}(\xi) d\xi$$

$$= \frac{1}{(2\pi)^{2n}} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \left(\int_{\mathbb{R}^n} e^{ix \cdot \eta} \sigma(x, \xi) \widehat{a}(\eta) d\eta \right) \widehat{f}(\xi) d\xi,$$

we have

(4.2)

$$[\sigma(X, D), a]f(x) = C_n \int_{\mathbb{R}^n} e^{ix \cdot \xi} \left(\int_{\mathbb{R}^n} e^{ix \cdot \eta} (\sigma(x, \xi + \eta) - \sigma(x, \xi)) \widehat{a}(\eta) d\eta \right) \widehat{f}(\xi) d\xi$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$, where $C_n = (2\pi)^{-2n}$. We decompose σ and a as follows:

$$\sigma(x, \xi) = \sum_{Q, Q' \in \mathcal{Q}} \sigma_{Q, Q'}(x, \xi) \quad \text{and} \quad a(x) = \sum_{j=0}^{\infty} \varphi_j(D) a(x),$$

where $\sigma_{Q, Q'}(x, \xi) = \psi_Q(D_x) \psi_{Q'}(D_\xi) \sigma(x, \xi)$ and $\{\varphi_j\}_{j \geq 0}$ is as in (2.11). Then

$$(4.3) \quad [\sigma(X, D), a] = \sum_{Q, Q' \in \mathcal{Q}} [\sigma_{Q, Q'}(X, D), \varphi_0(D) a] + \sum_{j=1}^{\infty} [\sigma(X, D), \varphi_j(D) a].$$

Let us consider the first sum of the right hand side of (4.3). Note that $\sigma_{Q, Q'} \in C^\infty(\mathbb{R}^n \times \mathbb{R}^n)$. By (4.2) and Taylor's formula, we have

$$\begin{aligned} & [\sigma_{Q, Q'}(X, D), \varphi_0(D) a]f(x) \\ &= C_n \int_{\mathbb{R}^n} e^{ix \cdot \xi} \left\{ \int_{\mathbb{R}^n} e^{ix \cdot \eta} \left(\sum_{k=1}^n \eta_k \int_0^1 \partial_{\xi_k} \sigma_{Q, Q'}(x, \xi + t\eta) dt \right) \varphi_0(\eta) \widehat{a}(\eta) d\eta \right\} \widehat{f}(\xi) d\xi \\ &= \frac{C_n}{i} \int_{\mathbb{R}^n} e^{ix \cdot \xi} \left\{ \sum_{k=1}^n \int_0^1 \left(\int_{\mathbb{R}^n} e^{ix \cdot \eta} \partial_{\xi_k} \sigma_{Q, Q'}(x, \xi + t\eta) \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \right) dt \right\} \widehat{f}(\xi) d\xi \end{aligned}$$

for all $f \in \mathcal{S}(\mathbb{R}^n)$, where $\eta = (\eta_1, \dots, \eta_n) \in \mathbb{R}^n$. Then, by Theorem 1.1,

$$(4.4) \quad \begin{aligned} & \|[\sigma_{Q, Q'}(X, D), \varphi_0(D) a]\|_{\mathcal{I}_1} \\ & \leq C \sum_{k=1}^n \int_0^1 \left\| \int_{\mathbb{R}^n} e^{ix \cdot \eta} (\partial_{\xi_k} \sigma_{Q, Q'})(x, \xi + t\eta) \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \right\|_{M_{\alpha n/2, \alpha}^{1,1}} dt \end{aligned}$$

for all $Q, Q' \in \mathcal{Q}$. Set

$$\tau_{Q, Q'}^{k, t}(x, \xi) = \int_{\mathbb{R}^n} e^{ix \cdot \eta} (\partial_{\xi_k} \sigma_{Q, Q'})(x, \xi + t\eta) \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta.$$

Recall that $\text{supp } \psi_Q \subset Q$ (see the definition of BAPU) and $\text{supp } \varphi_0 \subset \{|\eta| \leq 2\}$. Since

$$\begin{aligned} \mathcal{F}_{x \rightarrow x'}[\tau_{Q, Q'}^{k, t}(x, \xi)] &= \int_{\mathbb{R}^n} \mathcal{F}_{x \rightarrow x'}[e^{ix \cdot \eta} (\partial_{\xi_k} \sigma_{Q, Q'})(x, \xi + t\eta)] \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \\ &= \int_{\mathbb{R}^n} \psi_Q(x' - \eta) \mathcal{F}_1[\partial_{\xi_k} \psi_{Q'}(D_\xi) \sigma](x' - \eta, \xi + t\eta) \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \end{aligned}$$

and

$$\begin{aligned} \mathcal{F}_{\xi \rightarrow \xi'}[\tau_{Q, Q'}^{k, t}(x, \xi)] &= \int_{\mathbb{R}^n} e^{ix \cdot \eta} \mathcal{F}_{\xi \rightarrow \xi'}[(\partial_{\xi_k} \sigma_{Q, Q'})(x, \xi + t\eta)] \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \\ &= \int_{\mathbb{R}^n} e^{i(x + t\xi') \cdot \eta} (i\xi'_k) \psi_{Q'}(\xi') \mathcal{F}_2[\psi_Q(D_x) \sigma](x, \xi') \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta, \end{aligned}$$

we see that

$$\text{supp } \mathcal{F}_{x \rightarrow x'}[\tau_{Q, Q'}^{k, t}(x, \xi)] \subset \{x' \in \mathbb{R}^n : x' \in Q + \overline{B(0, 2)}\},$$

$$\text{supp } \mathcal{F}_{\xi \rightarrow \xi'}[\tau_{Q,Q'}^{k,t}(x, \xi)] \subset \{\xi' \in \mathbb{R}^n : \xi' \in Q'\}.$$

Then, by (2.9), Lemma 2.1 and $\sup_{Q \in \mathcal{Q}} \|\mathcal{F}^{-1}\psi_Q\|_{L^1} < \infty$, we have

$$\begin{aligned} \left\| \tau_{Q,Q'}^{k,t} \right\|_{M_{\alpha n/2, \alpha}^{1,1}} &= \sum_{\substack{\tilde{Q} \in \mathcal{Q} \\ \tilde{Q} \cap (Q + \overline{B(0,2)}) \neq \emptyset}} \sum_{\substack{\tilde{Q}' \in \mathcal{Q} \\ \tilde{Q}' \cap Q' \neq \emptyset}} \langle x_{\tilde{Q}} \rangle^{\alpha n/2} \langle \xi_{\tilde{Q}'} \rangle^{\alpha n/2} \\ &\quad \times \left\| \psi_{\tilde{Q}}(D_x) \psi_{\tilde{Q}'}(D_\xi) \tau_{Q,Q'}^{k,t} \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &\leq C \langle x_Q \rangle^{\alpha n/2} \langle \xi_{Q'} \rangle^{\alpha n/2} \left\| \tau_{Q,Q'}^{k,t} \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)}. \end{aligned} \quad (4.5)$$

Let $\gamma \in \mathcal{S}(\mathbb{R}^n)$ be such that $|\gamma| \geq 1$ on $\{|\xi| \leq 4\}$ and $\text{supp } \widehat{\gamma} \subset \{|x| < 1\}$ (for the existence of such a function, see the proof of [10, Theorem 2.6]). Since $\varphi_0 = \varphi_0 \gamma / \gamma = \gamma (\varphi_0 / \gamma)$, we can write $\varphi_0 = \gamma \Phi$, where $\Phi = \varphi_0 / \gamma \in \mathcal{S}(\mathbb{R}^n)$. Then

$$\begin{aligned} \tau_{Q,Q'}^{k,t}(x, \xi) &= \int_{\mathbb{R}^n} e^{ix \cdot \eta} (\partial_{\xi_k} \sigma_{Q,Q'})(x, \xi + t\eta) \varphi_0(\eta) \widehat{\partial_k a}(\eta) d\eta \\ &= \int_{\mathbb{R}^n} e^{ix \cdot \eta} (\partial_{\xi_k} \sigma_{Q,Q'})(x, \xi + t\eta) \gamma(\eta) \widehat{\Phi(D)}(\widehat{\partial_k a})(\eta) d\eta. \end{aligned} \quad (4.6)$$

By (2.9), (2.12) and Lemma 4.3, we see that

$$\begin{aligned} \|\partial_{\xi_k} \sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} &\leq \sum_{\tilde{Q}' \in \mathcal{Q}} \|\partial_{\xi_k} (\psi_{\tilde{Q}'}(D_\xi) \sigma_{Q,Q'})\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &= \sum_{\tilde{Q}' \cap Q' \neq \emptyset} \int_{\mathbb{R}^n} \left\| [\partial_{\xi_k} (\mathcal{F}^{-1} \psi_{\tilde{Q}'})] * \sigma_{Q,Q'}(x, \cdot) \right\|_{L^1} dx \\ &\leq \sum_{\tilde{Q}' \cap Q' \neq \emptyset} \int_{\mathbb{R}^n} \|\partial_{\xi_k} (\mathcal{F}^{-1} \psi_{\tilde{Q}'})\|_{L^1} \|\sigma_{Q,Q'}(x, \cdot)\|_{L^1} dx \\ &\leq C \sum_{\tilde{Q}' \cap Q' \neq \emptyset} \langle \xi_{\tilde{Q}'} \rangle \|\sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \leq C n_0 \langle \xi_{Q'} \rangle \|\sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)}. \end{aligned} \quad (4.7)$$

On the other hand, by (2.5),

$$(4.8) \quad \text{supp } \mathcal{F}_{\xi \rightarrow \xi'} [\partial_{\xi_k} \sigma_{Q,Q'}(x, \xi)] \subset Q' \subset B(d_{Q'}, 2R_{Q'}) \quad \text{for all } x \in \mathbb{R}^n.$$

Noting $R_{Q'} \asymp |Q'|^{1/n} \asymp \langle \xi_{Q'} \rangle^\alpha$ (see (2.7)), we have by (2.8), (4.6), (4.7), (4.8) and Lemma 4.2

$$\begin{aligned} \left\| \tau_{Q,Q'}^{k,t} \right\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} &\leq C(1 + 2R_{Q'})^{n/2} \|\partial_{\xi_k} \sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|\Phi(D)(\partial_k a)\|_{L^\infty} \\ &\leq C R_{Q'}^{n/2} \langle \xi_{Q'} \rangle \|\sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|\mathcal{F}^{-1} \Phi\|_{L^1} \|\partial_k a\|_{L^\infty} \\ &\leq C \langle \xi_{Q'} \rangle^{\alpha n/2 + 1} \|\sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \|\nabla a\|_{L^\infty} \end{aligned} \quad (4.9)$$

for all $0 < t < 1$. Combining (4.4), (4.5) and (4.9), we have

$$\begin{aligned} &\sum_{Q, Q' \in \mathcal{Q}} \|[\sigma_{Q,Q'}(X, D), \varphi_0(D)a]\|_{\mathcal{I}_1} \\ &\leq C \|\nabla a\|_{L^\infty} \left(\sum_{Q, Q' \in \mathcal{Q}} \langle x_Q \rangle^{\alpha n/2} \langle \xi_{Q'} \rangle^{\alpha n + 1} \|\sigma_{Q,Q'}\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \right) \\ &= C \|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n + 1), \alpha}^{1,1}}. \end{aligned}$$

We next consider the second sum of the right hand side of (4.3). Since

$$\varphi_j(D)a(x) = \int_{\mathbb{R}^n} 2^{jn}(\mathcal{F}^{-1}\varphi)(2^j(x-y))(a(y)-a(x))dy$$

and a is a Lipschitz function, we have $\|\varphi_j(D)a\|_{L^\infty} \leq C2^{-j}\|\nabla a\|_{L^\infty}$ for all $j \geq 1$. Hence, by (2.2) and Theorem 1.1, we see that

$$\begin{aligned} & \sum_{j=1}^{\infty} \|[\sigma(X, D), \varphi_j(D)a]\|_{\mathcal{I}_1} \\ & \leq \sum_{j=1}^{\infty} (\|\sigma(X, D)(\varphi_j(D)a)\|_{\mathcal{I}_1} + \|(\varphi_j(D)a)\sigma(X, D)\|_{\mathcal{I}_1}) \\ & \leq 2 \sum_{j=1}^{\infty} \|\varphi_j(D)a\|_{\mathcal{L}(L^2)} \|\sigma(X, D)\|_{\mathcal{I}_1} = 2 \sum_{j=1}^{\infty} \|\varphi_j(D)a\|_{L^\infty} \|\sigma(X, D)\|_{\mathcal{I}_1} \\ & \leq C \sum_{j=1}^{\infty} 2^{-j} \|\nabla a\|_{L^\infty} \|\sigma\|_{M_{\alpha n/2, \alpha}^{1,1}} \leq C \|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}}. \end{aligned}$$

Consequently, we obtain Theorem 1.2 with $a \in \mathcal{S}(\mathbb{R}^n)$.

Finally, we consider the general case. Let a be a Lipschitz function on \mathbb{R}^n . Since $[\sigma(X, D), a] = 0$ if a is a constant function, we may assume $\|\nabla a\|_{L^\infty} \neq 0$. Then, by Lemma 4.1, we have

$$(4.10) \quad \langle [\sigma(X, D), a]f, g \rangle = \lim_{\epsilon \rightarrow 0} \langle [\sigma(X, D), a_\epsilon]f, g \rangle \quad \text{for all } f, g \in \mathcal{S}(\mathbb{R}^n),$$

where $\{a_\epsilon\}_{0 < \epsilon < \epsilon(a)} \subset \mathcal{S}(\mathbb{R}^n)$ satisfies $\|\nabla a_\epsilon\|_{L^\infty} \leq C\|\nabla a\|_{L^\infty}$ for all $0 < \epsilon < \epsilon(a)$. By (2.1) and Theorem 1.2 with $a \in \mathcal{S}(\mathbb{R}^n)$,

$$(4.11) \quad \begin{aligned} & \|[\sigma(X, D), a_\epsilon]\|_{\mathcal{L}(L^2)} \leq \|[\sigma(X, D), a_\epsilon]\|_{\mathcal{I}_1} \\ & \leq C\|\nabla a_\epsilon\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}} \leq C\|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}} \end{aligned}$$

for all $0 < \epsilon < \epsilon(a)$. Combining (4.10) and (4.11), we have

$$(4.12) \quad \|[\sigma(X, D), a]\|_{\mathcal{L}(L^2)} \leq C\|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}}.$$

Then, (4.10), (4.11) and (4.12) give

$$(4.13) \quad \langle [\sigma(X, D), a]f, g \rangle = \lim_{\epsilon \rightarrow 0} \langle [\sigma(X, D), a_\epsilon]f, g \rangle \quad \text{for all } f, g \in L^2(\mathbb{R}^n).$$

Let $\{f_j\}, \{g_j\}$ be orthonormal systems in $L^2(\mathbb{R}^n)$. It follows from (2.3), (4.11), (4.13) and Fatou's lemma that

$$\begin{aligned} \sum_{j=1}^{\infty} |\langle [\sigma(X, D), a]f_j, g_j \rangle| &= \sum_{j=1}^{\infty} \lim_{\epsilon \rightarrow 0} |\langle [\sigma(X, D), a_\epsilon]f_j, g_j \rangle| \\ &\leq \liminf_{\epsilon \rightarrow 0} \sum_{j=1}^{\infty} |\langle [\sigma(X, D), a_\epsilon]f_j, g_j \rangle| \\ &\leq \liminf_{\epsilon \rightarrow 0} \|[\sigma(X, D), a_\epsilon]\|_{\mathcal{I}_1} \leq C\|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}}. \end{aligned}$$

Therefore, taking the supremum over all orthonormal systems $\{f_j\}, \{g_j\}$ in $L^2(\mathbb{R}^n)$, we have by (2.3)

$$\|[\sigma(X, D), a]\|_{\mathcal{I}_1} \leq C\|\nabla a\|_{L^\infty} \|\sigma\|_{M_{(\alpha n/2, \alpha n+1), \alpha}^{1,1}}.$$

The proof is complete.

APPENDIX A. THE INCLUSION BETWEEN FUNCTION SPACES

We first consider the relation between $B_{(n/2, n/2)}^{1,1}$ and $M^{1,1}$. Let $1 \leq p, q \leq \infty$ and p' be the conjugate exponent of p (that is, $1/p + 1/p' = 1$). In [27, Theorem 3.1], Toft proved the inclusions

$$B_{n\nu_1(p,q)}^{p,q}(\mathbb{R}^n) \hookrightarrow M^{p,q}(\mathbb{R}^n) \hookrightarrow B_{n\nu_2(p,q)}^{p,q}(\mathbb{R}^n),$$

where

$$\begin{aligned}\nu_1(p, q) &= \max\{0, 1/q - \min(1/p, 1/p')\}, \\ \nu_2(p, q) &= \min\{0, 1/q - \max(1/p, 1/p')\}\end{aligned}$$

(see also Gröbner [11], Okoudjou [20]). Due to [26, Theorem 1.2], the optimality of the inclusion relation between Besov and modulation spaces is described in the following way:

Proposition A.1. *Let $1 \leq p, q \leq \infty$ and $s \in \mathbb{R}$. Then the following are true:*

- (1) *If $B_s^{p,q}(\mathbb{R}^n) \hookrightarrow M^{p,q}(\mathbb{R}^n)$, then $s \geq n\nu_1(p, q)$.*
- (2) *If $M^{p,q}(\mathbb{R}^n) \hookrightarrow B_s^{p,q}(\mathbb{R}^n)$, then $s \leq n\nu_2(p, q)$.*

In particular, we have the best inclusions

$$(A.1) \quad B_n^{1,1}(\mathbb{R}^n) \hookrightarrow M^{1,1}(\mathbb{R}^n) \hookrightarrow B_0^{1,1}(\mathbb{R}^n).$$

Hence, we see that $B_{n/2}^{1,1}(\mathbb{R}^n)$ and $M^{1,1}(\mathbb{R}^n)$ have no inclusion relation with each other, and $B_{(n/2, n/2)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n)$ and $M^{1,1}(\mathbb{R}^{2n})$ also have the same relation since $\|f \otimes g\|_{M_{(s_1, s_2), \alpha}^{1,1}} = \|f\|_{M_{s_1, \alpha}^{1,1}} \|g\|_{M_{s_2, \alpha}^{1,1}}$ and $M_{(0,0)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n) = M^{1,1}(\mathbb{R}^{2n})$ (see Section 2). We remark that the statement (2) was shown in a restricted case $1 \leq p, q < \infty$ in [26], but it is also true for the endpoint $p = \infty$ or $q = \infty$ (see [17, Appendix A]).

We next give remarks on the relation between $M^{1,1}$ and $L_s^2 \cap H^s$. Recall that the norms on $L_s^2(\mathbb{R}^{2n})$ and $H^s(\mathbb{R}^{2n})$ are defined by

$$\begin{aligned}\|\sigma\|_{L_s^2} &= \left(\int_{\mathbb{R}^{2n}} \langle x; \xi \rangle^{2s} |\sigma(x, \xi)|^2 dx d\xi \right)^{1/2}, \\ \|\sigma\|_{H^s} &= \left(\int_{\mathbb{R}^{2n}} \langle x; \xi \rangle^{2s} |\widehat{\sigma}(x, \xi)|^2 dx d\xi \right)^{1/2},\end{aligned}$$

where $\langle x; \xi \rangle = (1 + |x|^2 + |\xi|^2)^{1/2}$ and $x, \xi \in \mathbb{R}^n$.

Proposition A.2. *The following are true:*

- (1) *If $s > 2n$, then $L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}) \hookrightarrow M^{1,1}(\mathbb{R}^{2n})$.*
- (2) *If $s \leq 2n$, then $L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}) \not\hookrightarrow M^{1,1}(\mathbb{R}^{2n})$.*
- (3) *If $s > n$, then $M^{1,1}(\mathbb{R}^{2n}) \not\hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$.*

Proof. We give the proof only for (3) because the assertions (1) and (2) were already proved in [14, Proposition 4.2]. Suppose, contrary to our claim, that $M^{1,1}(\mathbb{R}^{2n}) \hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ for $s > n$. Then, by (A.1),

$$B_{2n}^{1,1}(\mathbb{R}^{2n}) \hookrightarrow M^{1,1}(\mathbb{R}^{2n}) \hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}).$$

However, since $\langle x; \xi \rangle^{-2n-(s-n)/2} \in B_{2n}^{1,1}(\mathbb{R}^{2n})$ and $\langle x; \xi \rangle^{-2n-(s-n)/2} \notin L_s^2(\mathbb{R}^{2n})$ if $s > n$, this is a contradiction. \square

We finally consider the relation between $B_{(n/2, n/2)}^{1,1}$ and $L_s^2 \cap H^s$.

Proposition A.3. *The following are true:*

- (1) *If $s > 2n$, then $L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}) \hookrightarrow B_{(n/2, n/2)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n)$.*
- (2) *If $s > n$, then $B_{(n/2, n/2)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n) \not\hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$.*

Proof. Let $s > 2n$. By Schwarz's inequality,

$$\begin{aligned} \|\sigma\|_{B_{(n/2, n/2)}^{1,1}} &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} 2^{(j+k)n/2} \|\varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} 2^{(j+k)(n-s/2)/2} \|\langle x; \xi \rangle^{-s/2} \langle x; \xi \rangle^{s/2} 2^{(j+k)s/4} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^1} \\ &\leq \sum_{j=0}^{\infty} \sum_{k=0}^{\infty} 2^{(j+k)(n-s/2)/2} \|\langle x; \xi \rangle^{-s/2}\|_{L^2} \|\langle x; \xi \rangle^{s/2} 2^{(j+k)s/4} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2}, \end{aligned}$$

where $\{\varphi_j\}_{j \geq 0}$ is as in (2.11). Using $ab \leq (a^2 + b^2)/2$ for all $a, b \geq 0$, we have

$$\begin{aligned} &\|\langle x; \xi \rangle^{s/2} 2^{(j+k)s/4} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} \\ &\leq \frac{1}{2} \left(\|\langle x; \xi \rangle^s \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} + \|2^{(j+k)s/2} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} \right). \end{aligned}$$

Hence,

(A.2)

$$\begin{aligned} \|\sigma\|_{B_{(n/2, n/2)}^{1,1}} &\leq C \left(\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} 2^{(j+k)(n-s/2)/2} \|\langle x; \xi \rangle^s \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} \right) \\ &\quad + C \left(\sum_{j=0}^{\infty} \sum_{k=0}^{\infty} 2^{(j+k)(n-s/2)/2} \|2^{(j+k)s/2} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} \right) \end{aligned}$$

Let $\psi_j = \mathcal{F}^{-1} \varphi_j$, and we note that $\psi_j(x) = 2^{jn} \psi(2^j x)$ if $j \geq 1$, where $\psi = \mathcal{F}^{-1} \varphi$ and $\varphi \in \mathcal{S}(\mathbb{R}^n)$ is as in (2.11). Since

$$\begin{aligned} &|\langle x; \xi \rangle^s \varphi_j(D_x) \varphi_k(D_\xi) \sigma(x, \xi)| \\ &\leq C \int_{\mathbb{R}^{2n}} |\langle x - y; \xi - \eta \rangle^s \psi_j(x - y) \psi_k(\xi - \eta)| |\langle y; \eta \rangle^s \sigma(y, \eta)| dy d\eta \end{aligned}$$

and $\langle y; \eta \rangle^s \leq \langle 2^j y; 2^k \eta \rangle^s$, we have by Young's inequality

$$\begin{aligned} &\|\langle x; \xi \rangle^s \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} \\ (A.3) \quad &\leq C \left(\int_{\mathbb{R}^{2n}} |\langle 2^j y; 2^k \eta \rangle^s \psi_j(y) \psi_k(\eta)| dy d\eta \right) \|\langle x; \xi \rangle^s \sigma\|_{L^2} \leq C \|\langle x; \xi \rangle^s \sigma\|_{L^2} \end{aligned}$$

for all $j, k \geq 0$. On the other hand, since $2^{(j+k)s/2} \leq C \langle x; \xi \rangle^s$ for all $(x, \xi) \in \text{supp } \varphi_j \times \text{supp } \varphi_k$, we have

$$\begin{aligned} (A.4) \quad &\|2^{(j+k)s/2} \varphi_j(D_x) \varphi_k(D_\xi) \sigma\|_{L^2} = (2\pi)^{-n} \|2^{(j+k)s/2} (\varphi_j \otimes \varphi_k) \widehat{\sigma}\|_{L^2} \\ &\leq C \|\langle x; \xi \rangle^s \widehat{\sigma}\|_{L^2} \end{aligned}$$

for all $j, k \geq 0$. Combining (A.2), (A.3) and (A.4), we obtain (1).

We next consider (2). Assume that

$$(A.5) \quad B_{(n/2, n/2)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n) \hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n}),$$

where $s > n$. We note that

$$(A.6) \quad B_{s_1+s_2}^{1,1}(\mathbb{R}^{2n}) \hookrightarrow B_{(s_1, s_2)}^{1,1}(\mathbb{R}^n \times \mathbb{R}^n)$$

if $s_1, s_2 > 0$ (see [25, Theorem 1.3.9]). In fact, since $\text{supp } \Phi_0 \subset \{(x, \xi) : (|x|^2 + |\xi|^2)^{1/2} \leq 2\} \subset \{(x, \xi) : |x| \leq 2, |\xi| \leq 2\}$ and $\text{supp } \Phi_j \subset \{(x, \xi) : 2^{j-1} \leq (|x|^2 + |\xi|^2)^{1/2} \leq 2^{j+1}\} \subset \{(x, \xi) : |x| \leq 2^{j+1}, |\xi| \leq 2^{j+1}\}$, where $\Phi_0, \Phi_j \in \mathcal{S}(\mathbb{R}^{2n})$ are as in (2.11) with $2n$ instead of n , we have

$$\begin{aligned} \|\sigma\|_{B_{(s_1, s_2)}^{1,1}} &= \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} 2^{ks_1 + \ell s_2} \|\varphi_k(D_x) \varphi_{\ell}(D_{\xi}) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &\leq \sum_{k=0}^{\infty} \sum_{\ell=0}^{\infty} \sum_{j=0}^{\infty} 2^{ks_1 + \ell s_2} \|\varphi_k(D_x) \varphi_{\ell}(D_{\xi}) \Phi_j(D_{x, \xi}) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &= \sum_{j=0}^{\infty} \sum_{k=0}^{j+1} \sum_{\ell=0}^{j+1} 2^{ks_1 + \ell s_2} \|\varphi_k(D_x) \varphi_{\ell}(D_{\xi}) (\Phi_j(D_{x, \xi}) \sigma)\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \\ &\leq C \sum_{j=0}^{\infty} \|\Phi_j(D_{x, \xi}) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} \left(\sum_{k=0}^{j+1} 2^{ks_1} \right) \left(\sum_{\ell=0}^{j+1} 2^{\ell s_2} \right) \\ &\leq C \sum_{j=0}^{\infty} 2^{j(s_1 + s_2)} \|\Phi_j(D_{x, \xi}) \sigma\|_{L^1(\mathbb{R}^n \times \mathbb{R}^n)} = C \|\sigma\|_{B_{s_1+s_2}^{1,1}}. \end{aligned}$$

Then, it follows from (A.5) and (A.6) that $B_n^{1,1}(\mathbb{R}^{2n}) \hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$. However, this contradicts the fact that $B_{2n}^{1,1}(\mathbb{R}^{2n}) \not\hookrightarrow L_s^2(\mathbb{R}^{2n}) \cap H^s(\mathbb{R}^{2n})$ (see the proof of Proposition A.2). \square

APPENDIX B. PROOFS OF LEMMAS 2.1 AND 4.1

Proof of Lemma 2.1. Assume that $(Q + B(0, R)) \cap Q' \neq \emptyset$, where $Q, Q' \in \mathcal{Q}$.

We consider the first part. Let $\xi_{Q, Q'} \in (Q + B(0, R)) \cap Q'$. Since $\xi_{Q, Q'} = \widetilde{\xi_Q} + \xi$ for some $\widetilde{\xi_Q} \in Q$ and $\xi \in B(0, R)$, we see that $\langle \xi_{Q, Q'} \rangle \asymp \langle \widetilde{\xi_Q} \rangle$. Hence, by (2.9), $\langle \xi_Q \rangle \asymp \langle \widetilde{\xi_Q} \rangle \asymp \langle \xi_{Q, Q'} \rangle$, where $\xi_Q \in Q$. Similarly, $\langle \xi_{Q'} \rangle \asymp \langle \xi_{Q, Q'} \rangle$, where $\xi_{Q'} \in Q'$.

We next consider the second part. It follows from the first part that $|Q| \asymp \langle \xi_Q \rangle^{\alpha n} \asymp \langle \xi_{Q'} \rangle^{\alpha n} \asymp |Q'|$, and consequently

$$(B.1) \quad |Q| \asymp |Q'| \quad \text{if } (Q + B(0, R)) \cap Q' \neq \emptyset.$$

Let $B(c_Q, r_Q/2) \subset Q \subset B(d_Q, 2R_Q)$ and $B(c_{Q'}, r_{Q'}/2) \subset Q' \subset B(d_{Q'}, 2R_{Q'})$, where $Q, Q' \in \mathcal{Q}$ (see (2.5)). By (2.7) and (B.1), we see that $R_Q \asymp R_{Q'}$. Then, by (2.8),

$$\begin{aligned} \emptyset \neq (Q + B(0, R)) \cap Q' &\subset (B(d_Q, 2R_Q) + B(0, R)) \cap B(d_{Q'}, 2R_{Q'}) \\ &= B(d_Q, 2R_Q + R) \cap B(d_{Q'}, 2R_{Q'}) \subset B(d_Q, (2 + \kappa_2^{-1}R)R_Q) \cap B(d_{Q'}, 2R_{Q'}). \end{aligned}$$

Combining $B(d_Q, (2 + \kappa_2^{-1}R)R_Q) \cap B(d_{Q'}, 2R_{Q'}) \neq \emptyset$ and $R_Q \asymp R_{Q'}$, we obtain that $B(d_{Q'}, 2R_{Q'}) \subset B(d_Q, \kappa_3 R_Q)$ for some constant $\kappa_3 \geq 2$ independent of Q, Q' .

Hence, since $c_Q \in B(d_Q, \kappa_3 R_Q)$ and $r_Q \asymp R_Q$, if $(Q + B(0, R)) \cap Q' \neq \emptyset$ then

$$(B.2) \quad Q' \subset B(d_{Q'}, 2R_{Q'}) \subset B(d_Q, \kappa_3 R_Q) \subset B(c_Q, \kappa_4 r_Q),$$

where κ_4 is independent of $Q, Q' \in \mathcal{Q}$. Let $\mathcal{Q}_i, i = 1, \dots, n_0$, be subsets of \mathcal{Q} such that $\mathcal{Q} = \bigcup_{i=1}^{n_0} \mathcal{Q}_i$ and the elements of \mathcal{Q}_i are pairwise disjoint (see [1, Lemma B.1]). Set $A_Q = \{Q' \in \mathcal{Q} : (Q + B(0, R)) \cap Q' \neq \emptyset\}$. By (B.2), we have

$$\sum_{Q' \in A_Q \cap \mathcal{Q}_i} |Q'| \leq |B(c_Q, \kappa_4 r_Q)| = (2\kappa_4)^n |B(c_Q, r_Q/2)| \leq (2\kappa_4)^n |Q|$$

for all $1 \leq i \leq n_0$. Therefore, by (B.1), we see that

$$(\#A_Q)|Q| \leq \sum_{i=1}^{n_0} \sum_{Q' \in A_Q \cap \mathcal{Q}_i} (\kappa_5 |Q'|) \leq \kappa_5 \sum_{i=1}^{n_0} (2\kappa_4)^n |Q| = n_0 (2\kappa_4)^n \kappa_5 |Q|,$$

that is, $\#A_Q \leq n_0 (2\kappa_4)^n \kappa_5$. The proof is complete.

Proof of Lemma 4.1. Let $\varphi \in \mathcal{S}(\mathbb{R}^n)$ be such that $\varphi(0) = 1$, $\int_{\mathbb{R}^n} \varphi(x) dx = 1$ and $\text{supp } \varphi \subset \{x \in \mathbb{R}^n : |x| \leq 1\}$. If we set $a_\epsilon(x) = \varphi(\epsilon x)(\varphi_\epsilon * a)(x)$, then $\{a_\epsilon\}_{0 < \epsilon < \epsilon(a)} \subset \mathcal{S}(\mathbb{R}^n)$ satisfies (1) and (2), where $\varphi_\epsilon(x) = \epsilon^{-n} \varphi(x/\epsilon)$ and $\epsilon(a)$ will be chosen in the below.

We first consider (2). Since $|a(x) - a(y)| \leq \|\nabla a\|_{L^\infty} |x - y|$ for all $x, y \in \mathbb{R}^n$, we see that

$$\begin{aligned} |\partial_i(a_\epsilon(x))| &\leq \epsilon |(\partial_i \varphi)(\epsilon x) \varphi_\epsilon * a(x)| + |\varphi(\epsilon x) \varphi_\epsilon * (\partial_i a)(x)| \\ &\leq \epsilon |(\partial_i \varphi)(\epsilon x) (\varphi_\epsilon * a(x) - a(0))| + \epsilon |(\partial_i \varphi)(\epsilon x) a(0)| + \|\varphi\|_{L^1} \|\varphi\|_{L^\infty} \|\nabla a\|_{L^\infty} \\ &\leq \epsilon |(\nabla \varphi)(\epsilon x)| \int_{\mathbb{R}^n} \|\nabla a\|_{L^\infty} (1 + |x|)(1 + \epsilon |y|) |\varphi(y)| dy \\ &\quad + \epsilon |a(0)| \|\nabla \varphi\|_{L^\infty} + \|\varphi\|_{L^1} \|\varphi\|_{L^\infty} \|\nabla a\|_{L^\infty} \\ &\leq C_\varphi^1 C_\varphi^2 \|\nabla a\|_{L^\infty} + \epsilon |a(0)| \|\nabla \varphi\|_{L^\infty} + \|\varphi\|_{L^1} \|\varphi\|_{L^\infty} \|\nabla a\|_{L^\infty} \end{aligned}$$

for all $0 < \epsilon < 1$, where $C_\varphi^1 = \sup_{x \in \mathbb{R}^n} (1 + |x|) |\nabla \varphi(x)|$ and $C_\varphi^2 = \int_{\mathbb{R}^n} (1 + |y|) |\varphi(y)| dy$. Hence, we obtain (2) with $\epsilon(a) = \min\{\|\nabla a\|_{L^\infty}/|a(0)|, 1\}$ if $a(0) \neq 0$, and $\epsilon(a) = 1$ if $a(0) = 0$.

We next consider (1). Since a is continuous and $|a(x)| \leq C(1 + |x|)$ for all $x \in \mathbb{R}^n$, we see that $\lim_{\epsilon \rightarrow 0} a_\epsilon(x) = a(x)$ for all $x \in \mathbb{R}^n$, and $|a_\epsilon(x)| \leq C \|\varphi\|_{L^\infty} C_\varphi^2 (1 + |x|)$ for all $0 < \epsilon < \epsilon(a)$ and $x \in \mathbb{R}^n$. Hence, by the Lebesgue dominated convergence theorem, we have that $\lim_{\epsilon \rightarrow 0} \langle a_\epsilon T f, g \rangle = \langle a T f, g \rangle$ for all $f, g \in \mathcal{S}(\mathbb{R}^n)$, and $a_\epsilon f \rightarrow a f$ in $L^2(\mathbb{R}^n)$ as $\epsilon \rightarrow 0$ for all $f \in \mathcal{S}(\mathbb{R}^n)$, and consequently $T(a_\epsilon f) \rightarrow T(a f)$ in $L^2(\mathbb{R}^n)$ as $\epsilon \rightarrow 0$ for all $f \in \mathcal{S}(\mathbb{R}^n)$. The proof is complete.

REFERENCES

- [1] L. Borup and M. Nielsen, Banach frames for multivariate α -modulation spaces, J. Math. Anal. Appl. 321 (2006), 880-895.
- [2] L. Borup and M. Nielsen, Boundedness for pseudodifferential operators on multivariate α -modulation spaces, Ark. Mat. 44 (2006), 241-259.
- [3] A.P. Calderón, Commutators of singular integral operators, Proc. Nat. Acad. Sci. U.S.A. 53 (1965), 1092-1099.
- [4] R.R. Coifman and Y. Meyer, Commutateurs d'intégrales singulières et opérateurs multilinéaires, Ann. Inst. Fourier (Grenoble) 28 (1978), 177-202.
- [5] E. Cordero and K. Gröchenig, Time-frequency analysis of localization operators, J. Funct. Anal. 205 (2003), 107-131.

- [6] I. Daubechies, On the distributions corresponding to bounded operators in the Weyl quantization, *Comm. Math. Phys.* 75 (1980), 229-238.
- [7] H.G. Feichtinger and P. Gröbner, Banach spaces of distributions defined by decomposition methods, I, *Math. Nachr.* 123 (1985), 97-120.
- [8] C. Fernández and A. Galbis, Compactness of time-frequency localization operators on $L^2(\mathbb{R}^d)$, *J. Funct. Anal.* 233 (2006), 335-350.
- [9] M. Fornasier, Banach frames for α -modulation spaces *Appl. Comput. Harmon. Anal.* 22 (2007), 157-175.
- [10] M. Frazier and B. Jawerth, Decomposition of Besov spaces, *Indiana Univ. Math. J.* 34 (1985), 777-799.
- [11] P. Gröbner, Banachräume Glatter Funktionen und Zerlegungsmethoden, Thesis, University of Vienna, 1983.
- [12] K. Gröchenig, An uncertainty principle related to the Poisson summation formula, *Studia Math.* (1996), 87-104.
- [13] K. Gröchenig, A pedestrian approach to pseudodifferential operators, in: C. Heil (Ed.), *Harmonic Analysis and Applications*, Birkhäuser, Boston, 2006, pp. 139-169.
- [14] K. Gröchenig and C. Heil, Modulation spaces and pseudodifferential operators, *Integral Equations Operator Theory* 34 (1999), 439-457.
- [15] C. Heil, J. Ramanathan and P. Topiwara, Singular values of compact pseudodifferential operators, *J. Funct. Anal.* 150 (1997), 426-452.
- [16] L. Hörmander, The Weyl calculus of pseudodifferential operators, *Comm. Pure Appl. Math.* 32 (1979), 360-444.
- [17] M. Kobayashi, M. Sugimoto and N. Tomita, On the L^2 -boundedness of pseudo-differential operators and their commutators with symbols in α -modulation spaces, submitted, Available: arXiv:0707.0391.
- [18] D. Labate, Pseudodifferential operators on modulation spaces, *J. Math. Anal. Appl.* (2001), 242-255.
- [19] J. Marschall, Pseudo-differential operators with nonregular symbols of the class $S_{\rho,\delta}^m$, *Comm. Partial Differential Equations* 12 (1987), 921-965.
- [20] K. Okoudjou, Embedding of some classical Banach spaces into modulation spaces, *Proc. Amer. Math. Soc.* 132 (2004), 1639-1647.
- [21] J.C.T. Pool, Mathematical aspects of the Weyl correspondence, *J. Math. Phys.* 7 (1966), 66-76.
- [22] B. Simon, *Trace Ideals and their Applications*, Mathematical Surveys and Monographs 120, Amer. Math. Soc., second edition, 2005.
- [23] J. Sjöstrand, An algebra of pseudodifferential operators, *Math. Res. Lett.* 1 (1994), 185-192.
- [24] E.M. Stein, *Singular Integrals and Differentiability Properties of Functions*, Princeton University Press, Princeton, 1970.
- [25] M. Sugimoto, L^p -boundedness of pseudo-differential operators satisfying Besov estimates I, *J. Math. Soc. Japan* 40 (1988), 105-122.
- [26] M. Sugimoto and N. Tomita, The dilation property of modulation spaces and their inclusion relation with Besov spaces, *J. Funct. Anal.* 248 (2007), 79-106.
- [27] J. Toft, Continuity properties for modulation spaces, with applications to pseudo-differential calculus, I, *J. Funct. Anal.* 207 (2004), 399-429.
- [28] J. Toft, Continuity and Schatten properties for pseudo-differential operators on modulation spaces, *Operator Theory: Advances and Applications* Vol. 172, pp. 173-206.
- [29] J. Toft, Continuity and Schatten properties for Toeplitz operators on modulation spaces, *Operator Theory: Advances and Applications* Vol. 172, pp. 313-328.
- [30] H. Triebel, *Theory of Function spaces*, Birkhäuser, Basel-Boston-Stuttgart, 1983.
- [31] K. Zhu, *Operator Theory in Function Spaces*, Mathematical Surveys and Monographs 138, Amer. Math. Soc., second edition, 2007.

MITSURU SUGIMOTO, DEPARTMENT OF MATHEMATICS, GRADUATE SCHOOL OF SCIENCE, OSAKA
UNIVERSITY, TOYONAKA, OSAKA 560-0043, JAPAN

E-mail address: `sugimoto@math.sci.osaka-u.ac.jp`

NAOHITO TOMITA, DEPARTMENT OF MATHEMATICS, GRADUATE SCHOOL OF SCIENCE, OSAKA
UNIVERSITY, TOYONAKA, OSAKA 560-0043, JAPAN

E-mail address: `tomita@gaia.math.wani.osaka-u.ac.jp`